

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 15:23:02 ; Search time 8691 Seconds
(without alignments)

11503.997 Million cell updates/sec

Title: US-09-930-591-1

Perfect score: 2061

Sequence: 1 atggcgctatcacggccta.....atgaaatgaagagtgtga 2061

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2061	100.0	2061	6	CQ826998 Sequence
2	2061	100.0	2061	6	AX441176 Sequence
3	2061	100.0	2061	6	AX467113 Sequence
4	1854.8	90.0	9610	14	HEC278830
5	1846.4	89.6	8791	14	AY615798
6	1787	86.7	6299	6	AX164584
7	1786	86.7	5360	6	I06434
8	1786	86.7	5360	6	I09328
9	1786	86.7	6785	6	I06440
10	1786	86.7	7310	6	AR118696
11	1786	86.7	7310	6	I09331
12	1786	86.7	7310	14	HPCPOLYP
13	1786	86.7	9185	6	I08294
14	1786	86.7	9185	6	BD091382
15	1786	86.7	9379	6	AR166930
16	1786	86.7	9379	6	AR301300
17	1786	86.7	9401	6	AR176483
18	1786	86.7	9401	6	E66593
19	1786	86.7	9401	6	I71894

20	1786	86.7	9401	6	I81885	I81885 Sequence 9
21	1786	86.7	9401	6	BD080334	BD080334 Hepatitis
22	1786	86.7	9401	14	HPCLYPRE	M62321 Hepatitis C
23	1784.4	86.6	6785	6	AR118692	AR118692 Sequence
24	1784.4	86.6	6785	6	I09329	I09329 Sequence 10
25	1784.4	86.6	8316	6	AR118703	AR118703 Sequence
26	1784.4	86.6	8987	6	AR118728	AR118728 Sequence
27	1784.4	86.6	9185	6	AR118722	AR118722 Sequence
28	1784.4	86.6	9185	6	AR118723	AR118723 Sequence
29	1784.4	86.6	9379	6	AR118747	AR118747 Sequence
30	1781.2	86.4	2058	6	AR404933	AR404933 Sequence
31	1781.2	86.4	2058	6	AX395309	AX395309 Sequence
32	1781.2	86.4	5360	6	AR118686	AR118686 Sequence
33	1781.2	86.4	9424	14	AF511948	AF511948 Hepatitis
34	1781.2	86.4	9609	12	AF387805	AF387805 Synthetic
35	1781.2	86.4	9618	14	AF271632	AF271632 Hepatitis
36	1781.2	86.4	9646	12	AF387806	AF387806 Synthetic
37	1781.2	86.4	9693	12	AF387807	AF387807 Synthetic
38	1779.6	86.3	2058	6	AR408362	AR408362 Sequence
39	1779.6	86.3	2058	6	AX454818	AX454818 Sequence
40	1774.8	86.1	12980	6	AR110831	AR110831 Sequence
41	1774.8	86.1	12980	6	BD069985	BD069985 Functiona
42	1773.2	86.0	9609	12	AF387808	AF387808 Synthetic
43	1773.2	86.0	9646	6	AR110828	AR110828 Sequence
44	1773.2	86.0	9646	6	BD069982	BD069982 Functiona
45	1773.2	86.0	9646	14	AF009606	AF009606 Hepatitis

ALIGNMENTS

RESULT 1						
LOCUS	CQ826998	2061 bp	DNA	linear	PAT 29-JUN-2004	
DEFINITION	Sequence 1 from Patent WO2004048402.					
ACCESSION	CQ826998					
VERSION	CQ826998.1	GI:49455655				
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	synthetic construct					
REFERENCE	1					
AUTHORS	Sallberg,M.					
TITLE	A hepatitis c virus codon optimized non-structural ns3/4a fusion gene					
JOURNAL	Patent: WO 2004048402-A 1 10-JUN-2004;					
FEATURES	TRIPEP AB (SE)					
source	Location/Qualifiers					
	1..2061					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
	/note="Hepatitis C virus NS3/4A coding region"					

ORIGIN						
Query Match	100.0%;	Score 2061;	DB 6;	Length 2061;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 2061;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;					
Qy	1	ATGGCGCCTATCACGGCCTATGCCAGCAGACAAAGGGCCCTTTGGGATGCATATCACC	60			
Db	1	ATGGCGCCTATCACGGCCTATGCCAGCAGACAAAGGGCCCTTTGGGATGCATATCACC	60			
Qy	61	AGCTTACCGCGCGGACAAACACAGGTGAGGTTCAGATCGTGCACTGCT	120			
Db	61	AGCTTACCGCGCGGACAAACACAGGTGAGGTTCAGATCGTGCACTGCT	120			
Qy	121	GCCCAGACATTTCTTGGCAACCTGCATTAACGGGTGTGTGGACTGTCTACCATGGAGCC	180			
Db	121	GCCCAGACATTTCTTGGCAACCTGCATTAACGGGTGTGTGGACTGTCTACCATGGAGCC	180			
Qy	181	GGAAACAGGACCATTCGTCACCTAAGGTCCTGTTATCCAGATGTACCAATGTGGAC	240			
Db	181	GGAAACAGGACCATTCGTCACCTAAGGTCCTGTTATCCAGATGTACCAATGTGGAC	240			

```
Db 181 GGAACAAGGACCAATTGCGTCACCTAAGGGTCTGTGTATTCAGATGTATACCAATGTGGAC 240
QY 241 CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCGCTCAITTAACACCATGCACTTGC 300
Db 241 CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCGCTCAITTAACACCATGCACTTGC 300
QY 301 GGCTCCTCGGACCTTTACTCGGTCAACGAGGACGCCGATGTCATTCCTGTGTGCGCCGACGG 360
Db 301 GGCTCCTCGGACCTTTACTCGGTCAACGAGGACGCCGATGTCATTCCTGTGTGCGCCGACGG 360
QY 361 GGTGATGCGAGGCGAGCTGCTTTTCGCCCCGGCTATCTCTTACTTTGAAAGGCTCCTCG 420
Db 361 GGTGATGCGAGGCGAGCTGCTTTTCGCCCCGGCTATCTCTTACTTTGAAAGGCTCCTCG 420
QY 421 GGAGGCTCTGCTGTCGCCCGCAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGC 480
Db 421 GGAGGCTCTGCTGTCGCCCGCAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGC 480
QY 481 ACCCGTGGAGTGGCTAAGGGGTGGACTTCATCCCCGTAGAGACTTGAAGACAAACCATG 540
Db 481 ACCCGTGGAGTGGCTAAGGGGTGGACTTCATCCCCGTAGAGACTTGAAGACAAACCATG 540
QY 541 AGTTCGCCGCTGTTCTCAGACAACTCTCCCAACAGCAGTGCCTCCAGACTACCAAGTG 600
Db 541 AGTTCGCCGCTGTTCTCAGACAACTCTCCCAACAGCAGTGCCTCCAGACTACCAAGTG 600
QY 601 GCCACCTGCATGCTCCACCGCAGCGGTGAAGACACCAAGTCCCGCGGATACGCA 660
Db 601 GCCACCTGCATGCTCCACCGCAGCGGTGAAGACACCAAGTCCCGCGGATACGCA 660
QY 661 GCTCAGGGCTACAAGTGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGT 720
Db 661 GCTCAGGGCTACAAGTGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGT 720
QY 721 GCTTACATGTCGAAGGCCATGGGATGATCTTAACATCAGGACTGGGGTGAGGACAAAT 780
Db 721 GCTTACATGTCGAAGGCCATGGGATGATCTTAACATCAGGACTGGGGTGAGGACAAAT 780
QY 781 ACTACTGGCAGCCGATCAGTATTCACCTACGCAAGTTCCTTCCGACGCGCGGTGT 840
Db 781 ACTACTGGCAGCCGATCAGTATTCACCTACGCAAGTTCCTTCCGACGCGCGGTGT 840
QY 841 TCAGGGGGTCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAACATCC 900
Db 841 TCAGGGGGTCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAACATCC 900
QY 901 ATCTTGGGCAATGGCACTGCTTACCAAGAGAGACCGCGGGGGGAGACTGACTGTG 960
Db 901 ATCTTGGGCAATGGCACTGCTTACCAAGAGAGACCGCGGGGGGAGACTGACTGTG 960
QY 961 CTGCGCACCGCTACCCCTCGGGCTCGTCACTGTCGCCCATCTAACATCGAGGAGTT 1020
Db 961 CTGCGCACCGCTACCCCTCGGGCTCGTCACTGTCGCCCATCTAACATCGAGGAGTT 1020
QY 1021 GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAAT 1080
Db 1021 GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAAT 1080
QY 1081 AAGGGGGGAGACATCTACTTCTGCACTCAAGAGAGTGCACAGCTGCGCGCA 1140
Db 1081 AAGGGGGGAGACATCTACTTCTGCACTCAAGAGAGTGCACAGCTGCGCGCA 1140
QY 1141 AAAGTGTGCGGTGGCGCTCAATGCGGTCTTACTACCGGCGCTTGTATGTCCGTC 1200
Db 1141 AAAGTGTGCGGTGGCGCTCAATGCGGTCTTACTACCGGCGCTTGTATGTCCGTC 1200
QY 1201 ATCCGACAGTGGTGAAGTGTGCTGGCAACTGAGCCCTCATGCCGCTTTACC 1260
Db 1201 ATCCGACAGTGGTGAAGTGTGCTGGCAACTGAGCCCTCATGCCGCTTTACC 1260
QY 1261 GCGGACTTCGATTCGGGTAGACTGCAACAGTGTGTACCAGACAGTGCATTCAG 1320
Db 1261 GCGGACTTCGATTCGGGTAGACTGCAACAGTGTGTACCAGACAGTGCATTCAG 1320
```

```
QY 1321 CTTGACCCCTACCTTCAACATTTGAGACAATCAAGCTTCCCAGGATGCTGTCTCCCGTACT 1380
Db 1321 CTTGACCCCTACCTTCAACATTTGAGACAATCAAGCTTCCCAGGATGCTGTCTCCCGTACT 1380
QY 1381 CAACGTCGGGTAGGACTGCGAGAGGAAAGCAGGATCTACAGATTTGTGGCACCGGG 1440
Db 1381 CAACGTCGGGTAGGACTGCGAGAGGAAAGCAGGATCTACAGATTTGTGGCACCGGG 1440
QY 1441 GAGCGTCTTTCTGGCAATGTTTGAATCTGCTGCTCTGCGAGTGTCTATGACCGGGTTGT 1500
Db 1441 GAGCGTCTTTCTGGCAATGTTTGAATCTGCTGCTCTGCGAGTGTCTATGACCGGGTTGT 1500
QY 1501 GCTTGTATGAGCTTAGGCGCGCGAGACCAAGTTAGGCTACGAGCATACATGAACACC 1560
Db 1501 GCTTGTATGAGCTTAGGCGCGCGAGACCAAGTTAGGCTACGAGCATACATGAACACC 1560
QY 1561 CCGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCTCTTTACGGGTCTC 1620
Db 1561 CCGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCTCTTTACGGGTCTC 1620
QY 1621 ACCACATAGACGCCCATTTCTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
Db 1621 ACCACATAGACGCCCATTTCTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
QY 1681 CTGGTAGCGTACCAAGCCACCGTGTGCGTAGAGCTCAAGCCCCCTCCCGCTCGTGGAC 1740
Db 1681 CTGGTAGCGTACCAAGCCACCGTGTGCGTAGAGCTCAAGCCCCCTCCCGCTCGTGGAC 1740
QY 1741 CAGATGTGAAGTGTCTGATCCGTCTCAAGCCACCCCTCCATGGGCCAACACTCTGCTA 1800
Db 1741 CAGATGTGAAGTGTCTGATCCGTCTCAAGCCACCCCTCCATGGGCCAACACTCTGCTA 1800
QY 1801 TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTTGAGCACCAGTACCAAGTATATC 1860
Db 1801 TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTTGAGCACCAGTATATC 1860
QY 1861 ATGACATGTATCTCGGCTGACCTGAGGTGCTCAGAGTACCTGCTGCTGCTGGCGG 1920
Db 1861 ATGACATGTATCTCGGCTGACCTGAGGTGCTCAGAGTACCTGCTGCTGCTGGCGG 1920
QY 1921 GTTCTGGCTGCTTTGGCGCGCTATTGCTCTATCCAGAGTGTGCTCATAGTAGTAGG 1980
Db 1921 GTTCTGGCTGCTTTGGCGCGCTATTGCTCTATCCAGGCTGCTGCTCATAGTAGTAGG 1980
QY 1981 ATTGTCTTGTCCGAAAGCGGCAATCATACCCGACAGGAAAGTCTCTTACCGGAGTTC 2040
Db 1981 ATTGTCTTGTCCGAAAGCGGCAATCATACCCGACAGGAAAGTCTCTTACCGGAGTTC 2040
QY 2041 GATGAATGGAAGAGTGTCTGA 2061
Db 2041 GATGAATGGAAGAGTGTCTGA 2061
```

RESULT 2

```
AX441176
LOCUS AX441176 2061 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 16 from Patent WO0213855.
ACCESSION AX441176
VERSION AX441176.1 GI:21665758
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1
AUTHORS Sallberg, M. and Hultgren, C.
TITLE Vaccines containing ribavirin and methods of use thereof
JOURNAL Patent: WO 0213855-A 16 21-FEB-2002;
FEATURES
location/Qualifiers
1..2061
/organism="synthetic construct"
/mol_type="unassigned DNA"
```

/db_xref="taxon:32630" /note="Hepatitis C virus NS3/4A coding region"									
ORIGIN									
Query Match 100.0%; Score 2061; DB 6; Length 2061; Best Local Similarity 100.0%; Pred. No. 0; Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ATGCGCCTATACGGCCTATGCCAGCAGACAAAGGGGCCCTTTGGGATGCATAATCAC	60						
Db	1	ATGCGCCTATACGGCCTATGCCAGCAGACAAAGGGGCCCTTTGGGATGCATAATCAC	60						
Qy	61	AGCTTGACCGCGCGGAGAAAACACAGGTGAGGGTGAAGTTCAGATCGTGCATCGCT	120						
Db	61	AGCTTGACCGCGCGGAGAAAACACAGGTGAGGGTGAAGTTCAGATCGTGCATCGCT	120						
Qy	121	GCCAGACCTTTCTTGGCAACTGCAATTAACGGGTGTGGAGCTGTACCATGGAGCC	180						
Db	121	GCCAGACCTTTCTTGGCAACTGCAATTAACGGGTGTGGAGCTGTACCATGGAGCC	180						
Qy	181	GGAAACAGGACCAATGCGGTCACTTAAGGGTCTGTATCCAGATGTACACCAATGTGG	240						
Db	181	GGAAACAGGACCAATGCGGTCACTTAAGGGTCTGTATCCAGATGTACACCAATGTGG	240						
Qy	241	CAAGACCTGTAGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGC	300						
Db	241	CAAGACCTGTAGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGC	300						
Qy	301	GGCTCTCGGACCTTTACCTGGTCAAGGACAGCCGATCTATCTCTGTCGCGCAGCG	360						
Db	301	GGCTCTCGGACCTTTACCTGGTCAAGGACAGCCGATCTATCTCTGTCGCGCAGCG	360						
Qy	361	GGTGATGGCAGGGCAGCCTGCTTTCCGCCCGGCTATCTTACTTTGAAAGGCTCCTCG	420						
Db	361	GGTGATGGCAGGGCAGCCTGCTTTCCGCCCGGCTATCTTACTTTGAAAGGCTCCTCG	420						
Qy	421	GGAGGCCCTCTGTGTGCCCGCAGGACATGCGGTAGGCGATATTCAGAGCCGGGTATGC	480						
Db	421	GGAGGCCCTCTGTGTGCCCGCAGGACATGCGGTAGGCGATATTCAGAGCCGGGTATGC	480						
Qy	481	ACCGGTGGAGTGGCTAAGGGGTGGACTTATCCCGTGAAGAGCTTATAGAGCAACCATG	540						
Db	481	ACCGGTGGAGTGGCTAAGGGGTGGACTTATCCCGTGAAGAGCTTATAGAGCAACCATG	540						
Qy	541	AGGTCCCGGTGTCTCAGACAACTCTCCCGCAGCAGTCCCGCAGAGCTACCAAGTG	600						
Db	541	AGGTCCCGGTGTCTCAGACAACTCTCCCGCAGCAGTCCCGCAGAGCTACCAAGTG	600						
Qy	601	GCCACCTGCATGCTCCACCGGCGGTAAAGACACCAAGGTCCCGCGCATACGCA	660						
Db	601	GCCACCTGCATGCTCCACCGGCGGTAAAGACACCAAGGTCCCGCGCATACGCA	660						
Qy	661	GCTCAGGGCTACAGGTGCTGTGCTCAACCCCTCCGTTGCTGCCAATGGGCTTTGGT	720						
Db	661	GCTCAGGGCTACAGGTGCTGTGCTCAACCCCTCCGTTGCTGCCAATGGGCTTTGGT	720						
Qy	721	GCTTACATGTCCAAAGGCCCATGGATTGATCCTAAACATCAGGACTGGGGTGAGACAAT	780						
Db	721	GCTTACATGTCCAAAGGCCCATGGATTGATCCTAAACATCAGGACTGGGGTGAGACAAT	780						
Qy	781	ACTATGGCAGCCCGATCAAGTATTCACCTAGCGGAAGTTCCTTCCGACCGGGGTGT	840						
Db	781	ACTATGGCAGCCCGATCAAGTATTCACCTAGCGGAAGTTCCTTCCGACCGGGGTGT	840						
Qy	841	TCAGGGGTGCTTATGATAAATTTGTGACAGTGCCTCCAGCGATGCACATCC	900						
Db	841	TCAGGGGTGCTTATGATAAATTTGTGACAGTGCCTCCAGCGATGCACATCC	900						
Qy	901	ATCTTGGGCACTGGCACTGCTTGAACAAGCAGAGACCGGGGGGAGACTGACTGTG	960						
Db	901	ATCTTGGGCACTGGCACTGCTTGAACAAGCAGAGACCGGGGGGAGACTGACTGTG	960						
Qy	961	CTCGCCACCGCTACCCCTCCGGGCTCCGTCACCTGTGCCCCATCTTAACATCGAGGAGGTT	1020						

Db	961	CTCGCCACCGCTACCCCTCCGGGCTCCGTCACCTGTGCCCCATCTTAACATCGAGGAGTT	1020						
Qy	1021	GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAAT	1080						
Db	1021	GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAAT	1080						
Qy	1081	AAGGGGGGAGACATCTCATCTTTCCTCACTCAAAAGAAAGTGCAGAGCTCGCGCA	1140						
Db	1081	AAGGGGGGAGACATCTCATCTTTCCTCACTCAAAAGAAAGTGCAGAGCTCGCGCA	1140						
Qy	1141	AAACTGGTCCGCTGGCGTCAATGCGTGCTTACTACCGCGCTTGTATGTGTCGCTC	1200						
Db	1141	AAACTGGTCCGCTGGCGTCAATGCGTGCTTACTACCGCGCTTGTATGTGTCGCTC	1200						
Qy	1201	ATCCCGACCAAGTGTGACGTGTGCTGCGCAACTGACGCCCTCATGACCGGCTTTACC	1260						
Db	1201	ATCCCGACCAAGTGTGACGTGTGCTGCGCAACTGACGCCCTCATGACCGGCTTTACC	1260						
Qy	1261	GGCGACTTCGATTTCGGTGTAGACTGCAACACGCTGTGTCAACAGACAGTTCAGC	1320						
Db	1261	GGCGACTTCGATTTCGGTGTAGACTGCAACACGCTGTGTCAACAGACAGTTCAGC	1320						
Qy	1321	CTTGACCTTACCTTCACCATTTGAGCAATCAACGCTTCCCAGGATGCTCTCCGTA	1380						
Db	1321	CTTGACCTTACCTTCACCATTTGAGCAATCAACGCTTCCCAGGATGCTCTCCGTA	1380						
Qy	1381	CAACGTCCGGGTAGGACTGGCAGAGGAAGCAGGCAATCTACAGATTTGTGGACCGGG	1440						
Db	1381	CAACGTCCGGGTAGGACTGGCAGAGGAAGCAGGCAATCTACAGATTTGTGGACCGGG	1440						
Qy	1441	GAGCGTCTTCTCGGCAATTTTGAATCTGCTGTCTCTGCGAGTGTCTACGCGGTTGT	1500						
Db	1441	GAGCGTCTTCTCGGCAATTTTGAATCTGCTGTCTCTGCGAGTGTCTACGCGGTTGT	1500						
Qy	1501	GCTTGTATGAGCTTACGCGCGCGAGACCAAGTTAGGCTACGAGCATACATGAACACC	1560						
Db	1501	GCTTGTATGAGCTTACGCGCGCGAGACCAAGTTAGGCTACGAGCATACATGAACACC	1560						
Qy	1561	CCGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGCGCTTTTACGGGTCTC	1620						
Db	1561	CCGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGCGCTTTTACGGGTCTC	1620						
Qy	1621	ACCCACATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAAAACTTTCCCTAT	1680						
Db	1621	ACCCACATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAAAACTTTCCCTAT	1680						
Qy	1681	CTGGTAGCTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTTCCCGCTGTGGGAC	1740						
Db	1681	CTGGTAGCTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTTCCCGCTGTGGGAC	1740						
Qy	1741	CAGATGTGAAGTGTGATCCGCTCTCAAGCCACCTCCATGGGCCAACCTCTCTCTA	1800						
Db	1741	CAGATGTGAAGTGTGATCCGCTCTCAAGCCACCTCCATGGGCCAACCTCTCTCTA	1800						
Qy	1801	TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATC	1860						
Db	1801	TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATC	1860						
Qy	1861	ATGACATGTATGTCCGCTGACCTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	1920						
Db	1861	ATGACATGTATGTCCGCTGACCTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	1920						
Qy	1921	GTTCTGGCTGTTTGGCGCGTATTCCTATCCAGAGCTGCGTGGTTCATAGTAGTAGG	1980						
Db	1921	GTTCTGGCTGTTTGGCGCGTATTCCTATCCAGAGCTGCGTGGTTCATAGTAGTAGG	1980						
Qy	1981	ATTGTCTTGTCCGAAAGCGGCAATCATATCCGACAGGGAAGTCTCTCTACCGGAGTTC	2040						
Db	1981	ATTGTCTTGTCCGAAAGCGGCAATCATATCCGACAGGGAAGTCTCTCTACCGGAGTTC	2040						
Qy	2041	GATGAATGGAAGTGTCTGA 2061							

Db	2041 GATGAAATGGAAGTGTCTGA 2061				
RESULT 3					
AX467113					
LOCUS	AX467113	2061 bp	DNA	linear	PAT 17-JUL-2002
DEFINITION	Sequence 1 from Patent WO20214362.				
ACCESSION	AX467113				
VERSION	AX467113.1 GI:21900425				
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	other sequences; artificial sequences.				
REFERENCE	1. Sallberg, M.				
AUTHORS	A hepatitis c virus non-structural ns3/4a fusion gene				
TITLE	Patent: WO 0214362-A 1 21-FEB-2002;				
JOURNAL	TRIPBP AB (SE)				
FEATURES	Location/Qualifiers				
source	1..2061				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="Hepatitis C virus NS3/4A coding region"				
ORIGIN					
	Query Match	100.0%;	Score 2061;	DB 6;	Length 2061;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2061;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGCGCCTATCACGCGCTATGCCAGCAGACAAGGGGCTTTGGGATGCATATCAC	60		
Db	1	ATGCGCCTATCACGCGCTATGCCAGCAGACAAGGGGCTTTGGGATGCATATCAC	60		
Qy	61	AGCTTACCGCGCGGACAAAACACAGTGGAGGTGAGGTTGAGATCGTGCACCTGCT	120		
Db	61	AGCTTACCGCGCGGACAAAACACAGTGGAGGTGAGGTTGAGATCGTGCACCTGCT	120		
Qy	121	GCCACAGCTTTCTTGGCAACCTGTCATTAAACGGGGTGTGTGGACTGTCTACATGGAGCC	180		
Db	121	GCCACAGCTTTCTTGGCAACCTGTCATTAAACGGGGTGTGTGGACTGTCTACATGGAGCC	180		
Qy	181	GGNACGAAGACATTTGCGTCACTAAGGGTCTGTATTCAGATGTACACAAATGTGGAC	240		
Db	181	GGNACGAAGACATTTGCGTCACTAAGGGTCTGTATTCAGATGTACACAAATGTGGAC	240		
Qy	241	CAAGACTCTGCTAGGCTGGCGGCTCCCAAGGTGCGGCTCATTAACACCATGCACTTGC	300		
Db	241	CAAGACTCTGCTAGGCTGGCGGCTCCCAAGGTGCGGCTCATTAACACCATGCACTTGC	300		
Qy	301	GGCTCTCGGACCTTTAATCTGGTACAGAGGACACCGCATGTCAATTCCTGTGCGCCGACGG	360		
Db	301	GGCTCTCGGACCTTTAATCTGGTACAGAGGACACCGCATGTCAATTCCTGTGCGCCGACGG	360		
Qy	361	GGTGATGGCAGGGCAGCTGCTTTGCGCCCGGCTACTCTTACTTTGAAAGGCTCCTCG	420		
Db	361	GGTGATGGCAGGGCAGCTGCTTTGCGCCCGGCTACTCTTACTTTGAAAGGCTCCTCG	420		
Qy	421	GGAGGCCCTCTGCTGTGCCCCGACGACATGCCGTAGGCATATTCAGAGCCGGGTATGC	480		
Db	421	GGAGGCCCTCTGCTGTGCCCCGACGACATGCCGTAGGCATATTCAGAGCCGGGTATGC	480		
Qy	481	ACCCGTGGAGTGGCTTAAGCGGTGGACTTCAATCCCGTAGAGCTTTAGAGACAACATG	540		
Db	481	ACCCGTGGAGTGGCTTAAGCGGTGGACTTCAATCCCGTAGAGCTTTAGAGACAACATG	540		
Qy	541	AGGTCCCGGTGTTCTCAGACAACTCTCTCCCAACAGCAGTGCACAGACTTACCAAGTG	600		
Db	541	AGGTCCCGGTGTTCTCAGACAACTCTCTCCCAACAGCAGTGCACAGACTTACCAAGTG	600		
Qy	601	GCCACCTTCATGCTCCACCGCAGCGGTAAAGACACAAAGTCCCGCGCATACGCA	660		
Db	601	GCCACCTTCATGCTCCACCGCAGCGGTAAAGACACAAAGTCCCGCGCATACGCA	660		
Qy	661	GCTCAGGGCTACAAGTGTCTGCTCAACCCCTCGTTGCTGCTCAACAATGGGCTTTGGT	720		
Db	661	GCTCAGGGCTACAAGTGTCTGCTCAACCCCTCGTTGCTGCTCAACAATGGGCTTTGGT	720		
Qy	721	GCTTACATGTCCAAGGCCATGGGATTTGATCCTTAACATCAGGACTGGGGTGAAGACAATT	780		
Db	721	GCTTACATGTCCAAGGCCATGGGATTTGATCCTTAACATCAGGACTGGGGTGAAGACAATT	780		
Qy	781	ACTACTGGCAGCCCGATCACGTTATTCACCTACCGCAAGTTCCTTGGCGGCGGGTGT	840		
Db	781	ACTACTGGCAGCCCGATCACGTTATTCACCTACCGCAAGTTCCTTGGCGGCGGGTGT	840		
Qy	841	TCAGGGGTGCTTATGACATAATAATTTGTGACAGTGCCACTCCACGGATGCAACAATCC	900		
Db	841	TCAGGGGTGCTTATGACATAATAATTTGTGACAGTGCCACTCCACGGATGCAACAATCC	900		
Qy	901	ATCTTTGGGATTTGGCAGTCTCTTGACCAAGCAGACACCGGGGGGAGACATGACTGTG	960		
Db	901	ATCTTTGGGATTTGGCAGTCTCTTGACCAAGCAGACACCGGGGGGAGACATGACTGTG	960		
Qy	961	CTCGCACCGGTACCCCTCCGGGCTCCGTCACCTGTGCCCCATCTTAACATCGAGGAGTT	1020		
Db	961	CTCGCACCGGTACCCCTCCGGGCTCCGTCACCTGTGCCCCATCTTAACATCGAGGAGTT	1020		
Qy	1021	GCTCTGTCCACTACCGGAGAGATCCCTTTATGGCAAGGCTATTCCTTTGAAGCAATT	1080		
Db	1021	GCTCTGTCCACTACCGGAGAGATCCCTTTATGGCAAGGCTATTCCTTTGAAGCAATT	1080		
Qy	1081	AAGGGGGGAGACATCTCATCTTCTGGCCTCAAAAGAAAGTGGCAGCAGCTCGCGCA	1140		
Db	1081	AAGGGGGGAGACATCTCATCTTCTGGCCTCAAAAGAAAGTGGCAGCAGCTCGCGCA	1140		
Qy	1141	AAACTGTGCGGTGGGGCTCAATGCGTGGCTTACTACCGCGGCTTGTATGTCCGTC	1200		
Db	1141	AAACTGTGCGGTGGGGCTCAATGCGTGGCTTACTACCGCGGCTTGTATGTCCGTC	1200		
Qy	1201	ATCCGACACGTGTGAGTTGCTGCTGGCAACTGACGCCCTCATGACCGGCTTTACC	1260		
Db	1201	ATCCGACACGTGTGAGTTGCTGCTGGCAACTGACGCCCTCATGACCGGCTTTACC	1260		
Qy	1261	GGCGATTCGATTCGGTGTAGATGTCACACACGTTGTCTACCCAGACAGTTCGACTTCAGC	1320		
Db	1261	GGCGATTCGATTCGGTGTAGATGTCACACACGTTGTCTACCCAGACAGTTCGACTTCAGC	1320		
Qy	1321	CTTGACCTCACTTTCACCATTTGACAAATCACCGTTCCCGAGATGCTCTCCCGTACT	1380		
Db	1321	CTTGACCTCACTTTCACCATTTGACAAATCACCGTTCCCGAGATGCTCTCTCCCGTACT	1380		
Qy	1381	CAACGTGGGGTAGGACTGGCAGAGGGAAGCCAGGCATCTACAGATTTGTGGCACCAGGG	1440		
Db	1381	CAACGTGGGGTAGGACTGGCAGAGGGAAGCCAGGCATCTACAGATTTGTGGCACCAGGG	1440		
Qy	1441	GAGCGTCTTCTGCGATTTTGAATGCTGCTCTCTGCGAGTGTATGACGGGGTGT	1500		
Db	1441	GAGCGTCTTCTGCGATTTTGAATGCTGCTCTCTGCGAGTGTATGACGGGGTGT	1500		
Qy	1501	GCTTTGGTATGAGCTTTAGCCCGCGGAGACCAACAGTTAGGCTACGACATATGAACACC	1560		
Db	1501	GCTTTGGTATGAGCTTTAGCCCGCGGAGACCAACAGTTAGGCTACGACATATGAACACC	1560		
Qy	1561	CCGGGACTTCCCGTGTGCCAAGACCAATCTTTGAATTTTGGGAGGGCGCTCTTTACGGGTCTC	1620		
Db	1561	CCGGGACTTCCCGTGTGCCAAGACCAATCTTTGAATTTTGGGAGGGCGCTCTTTACGGGTCTC	1620		
Qy	1621	ACCACATAGACGCCACTTCTCTATCCACAGAAAACAGAGTGGGGAACCTTCCCTAT	1680		
Db	1621	ACCACATAGACGCCACTTCTCTATCCACAGAAAACAGAGTGGGGAACCTTCCCTAT	1680		
Qy	1681	CTGTGACGTTACCAAGCCACCGTGTGCGTGTAGAGCTCAAGCCCCCTCCCGCTGTGGGAC	1740		
Db	1681	CTGTGACGTTACCAAGCCACCGTGTGCGTGTAGAGCTCAAGCCCCCTCCCGCTGTGGGAC	1740		

/product="non-structural protein 5b"
/note="ORF10"

ORIGIN									
Query Match 90.0%; Score 1854.8; DB 14; Length 9610; Best Local Similarity 93.8%; Pred. No. 0; Matches 1931; Conservative 0; Mismatches 127; Indels 0; Gaps 0;									
Qy	2	TGGCGCCTATCACGGCCTATGCCAGAGACAAGGGCCCTTTGGGATGCATATACCA	61	Db	4378	TCGCGACCGCTACCCCTCCGGGCTCGTCACTGTGCCCCCATTTAAATCAAGAGGTTG	4437		
Db	3418	TGGGCGCCATCACGGGTACGCCAGACAGCAAGGGCCCTTTGGGATGCATATACCA	3477	Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTATATGGCAAGGCTATTCCTCTTGAAGCAATTA	1081		
Qy	62	GCTTGAACGGCCGGGACAAAACAGAGTGGAGGGTGAGGTTTCAGATCGTGTCAACTGCTG	121	Db	4438	CTCTGTCCACCAACCGGGGAGATTCTCTTTTATAGGCAAGGCTATTCCTCTCGAGGTATCA	4497		
Db	3478	GCCTGAACGGCCGGGACAAAACAGAGTGGAGGGTGAGGTTTCAGATCGTGTCAACTGCTG	3537	Qy	1082	AGGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAAGAGTGGCAGAGCTGCGCGCAA	1141		
Qy	122	CCGAGACTTCTTGGCAACCTGCATTAACGGGGGTGTGTGGACGTCTTACATGGAGCCG	181	Db	4498	AGGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAAGAGTGGCAGACCTCGCGCGCAA	4557		
Db	3538	CCGAGACTTCTTGGCAACCTGCATTAACGGGGGTGTGTGGACGTCTTACATGGAGCCG	3597	Qy	1142	AACCTGTGCGGTTGGCGGCTCAATGCGGTGCTTACTACCGCGCCCTTGTATGTTCGCTCA	1201		
Qy	182	GAAACAAGGACCATTCGCTCACTAAGGCTCTGTATTATCCAGATGTACACCAATGTGGACC	241	Db	4558	AACCTGTGCGGTTGGCGCATCAATGCGGTGCTTACTACCGCGCCCTTGTATGTTCGCTCA	4617		
Db	3598	GAAACAAGGACCATTCGCTCACTAAGGCTCTGTATTATCCAGATGTACACCAATGTGGACC	3657	Qy	1202	TCCCGACCAAGTGTGATGCTGTGCTGTGGCAACTGAACGCCCTCATGACCGGCTTTACCG	1261		
Qy	242	AAGACCTCGTAGGCTGCGCGCTCCCAAGGTGCGCGCTCAITTAACACCATGCACTTGGC	301	Db	4618	TCCCGACCAAGTGTGATGCTGTGCTGTGGCAACAGATGCTCTCATGACCGGCTTTACCG	4677		
Db	3658	AAGACCTCGTAGGCTGCGCGCTCCCAAGGTGCGCGCTCAITTAACACCATGCACTTGGC	3717	Qy	1262	GGGACTTCGATTCGGGTGATAGACTGCAACACCTGTGTCAACAGAGTGTGACTTCAGCC	1321		
Qy	302	GCTCCTCGGACCTTTACTGTGACGAGGACGCGCGATGCTATTCGTGCGCGGACGGG	361	Db	4678	GGGACTTCGATTCGGGTGATAGACTGCAACACATGTGTCAACAGAGTGTGACTTCAGCC	4737		
Db	3718	GCTCCTCGGACCTTTACTGTGACGAGGACGCGCGATGCTATTCGTGCGCGGCGGG	3777	Qy	1322	TTGACCTTACCTTACCATTTGAGACAATCAACGCTTCCCGCAGGATGCTGTCTCCCGTACTC	1381		
Qy	362	GTGATGTCAGGGGACGCTGCTTTTCGCGCCGCGCTATCTTCTTCTTGAAGGCTCTCTCGG	421	Db	4738	TGACCCCTTACTTTCACCATTTGAGACAAACACGCTTCCCGCAGGATGCTGTCTCCCGCACCC	4797		
Db	3778	GTGATGTCAGGGGACGCTGCTTTTCGCGCCGCGCTATCTTCTTCTTGAAGGCTCTCTCGG	3837	Qy	1382	AACGTGCGGGTAGGACTGCGCAGAGGAAAGCCAGGCATCTACAGATTTGTGGCAACCGGGG	1441		
Qy	422	GAGCCCTCTGCTGTGCGCGCAGGACATCCGCTAGGCAATATTACAGACCGCGGTATGCA	481	Db	4798	AACGTGCGGGCAGGACTGCGCAGGGGAAAGCCAGGCATCTACAGATTTGTGGCAACCGGGG	4857		
Db	3838	GGGGCCCACTGCTGTGCGCTGCGGACACGCGGTGGGCAATTTACGGCGCGCGGTATGTA	3897	Qy	1442	AGCGTCTTCTCGCATGTTTGAAGTCTGCTGCTCTGCTGCGAGTGTATGACGCGGGGTGTG	1501		
Qy	482	CCGTGAGGTGGCTTAAGCGCGTGGACTTCTATCCCGGTAGAGACTTAGAGACAACCATGA	541	Db	4858	AGCGCCCTTCCGGTATGTTGCACTCGCGCGTCTCTGCGAGTGTATGACGCGGGGTGTG	4917		
Db	3898	CCGAGAGTGGCTTAAGCGCGTGGAAATTTGCTCCCTGTAGAGAACCTTAGAGACAACCATGA	3957	Qy	1502	CTTGTGTATGAGCTTACGCCCGCGGAGACACAGTTTAGGCTACAGCATACATGAAACACCC	1561		
Qy	542	GGTCCCGGTGTTCTCAGACAATCTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTGG	601	Db	4918	CTTGTGTATGAGCTACACCCCGCGGAGACACAGTTTAGGCTACGAGCATACATGAAACACCC	4977		
Db	3958	GGTCCCGGTGTTCTCAGACAATCTCTCCCAACAGCAGTGCCTCCAGAGCTTCCAGGTGG	4017	Qy	1562	CGGGACTTCCCGGTGCGCAAGCAATCTTGAATTTTGGAGGGCGCTTTTACGGGTCTCA	1621		
Qy	602	CCACCTGCATGTCCCAACCGGAGCGGTAAAGACCAAGGTTCCTCCGCGCATACGCGAG	661	Db	4978	CGGGACTTCCCGGTGCGCAAGCAATCTTGAATTTTGGAGGGCGCTTCTTCACGGCTCTCA	5037		
Db	4018	CCACCTGCATGTCTCCCAACCGGAGCGGTAAAGACCAAGGTTCCTCCGCGCTTACGCGG	4077	Qy	1622	CCCAATAGACCCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC	1681		
Qy	662	CTCAGGGCTACAAGGTGCTGGTCTCAACCCCTCCGTGCTGCTGCAACAAATGGGCTTTGGTG	721	Db	5038	CCCAATAGACCGCTCACTTCTCTATCCAGACAAAGCAGAGTGGGAAACCTTCTTCTTACC	5097		
Db	4078	CTCAGGGTACAAGGTGCTGGTCTCAACCCCTCCGTGCTGCTGCAACAAATGGGCTTTGGTG	4137	Qy	1682	TGGTAGCGTACCAAGCCACCGTGTGCTGTAGAGCTCAAGCCCCCTCCCGCTCGTGGGACC	1741		
Qy	722	CTTACATGTCGAAGGCCATGGGATTCATCTTAACATCAGGACTGGGGTGAAGCAATTA	781	Db	5098	TGGTAGCGTACCAAGCCACCGTGTGCTGTAGAGCCCAAGCCCCCTCCCGCTCGTGGGACC	5157		
Db	4138	CCTATATGTCGAAGGCCATGGGATTCATCTTAACATCAGGACTGGGGTGAAGCAATTA	4197	Qy	1742	AGATGTGGAAGTGTGATCGCTCTCAAGCCCAACCTCCATGGGGCAACACCTCTGTCTAT	1801		
Qy	782	CTACTGGCAGCCGATCACTGATTCACCTTACGGCAAGTTCCTTTCGCGACGGCGGTGTT	841	Db	5158	AGATGTGGAAGTGTGATTCGCTCTCAAGCCCAACCTTCTCATGGGTCAACACCTCTGTCTAT	5217		
Db	4198	CCACCGGACGCCCATCACTACTTCCACTTACGGCAAGTTCCTTTCGCGACGGCGGTGCT	4257	Qy	1802	ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTTGACACCCAGTCAACCAAGTATATCA	1861		
Qy	842	CAGGGGTGCTTATGACATAATTAATTTGTGACGAGTGCCACTCCACGGATGCATATCCA	901	Db	5218	ACAGACTGGGCGCTGTTTCAAGATGAAGTCAACCTTAACACCCAGTCAACCAATATATCA	5277		
Db	4258	CAGGGGTGCTTATGACATAATTAATTTGTGACGAGTGCCACTCCACGGATGCATATCCA	4317	Qy	1862	TGACATGTATGTGCGCTGACCTTGGAGGTGCTCAGAGTACCTGGGTGCTCGTTGGCGCGG	1921		
Qy	902	TCTTGGGCAATGGCACTGTCTTTGACCAAGCAGAGACCGCGGGGCGAGCTGACTGTGCG	961	Db	5278	TGACATGTATGTGCGCTGATCTGGAGATCGTCACGAGTACCTGGGTGCTCGTTGGCGCGG	5337		
Db	4318	TCTTGGGCAATGGCACTGTCTTTGACCAAGCAGAGACCGCGGGGCGAGGCTGTTGTGCG	4377	Qy	1922	TTCTGGCTGCTTTGGCGCGGCTATTCCTTATCCAGGCTGCTGGTGTATAGTAGGTAGGA	1981		
Qy	962	TGCGCACCGCTACCCCTCCGGGCTCGTCACTGTGCCCCCATCTTAACATCGAGGAGGTTG	1021	Db	5338	TTCTGGCTGCTTTGGCGCGGCTACTGTCTATCCAGGCTGCTGGTGTATAGTAGGCAGGA	5397		
Db				Qy	1982	TTGTCTTCTTGGGAAAGCGCGCAATCATACCCGACAGGAAAGTCTCTTACCGGAGTTCG	2041		
Qy				Db	5398	TTGTCTTCTTGGGAAAGCGCGCAATCATACCCGACAGGAAAGTCTCTTACCGGAGTTCG	5457		
Db				Qy	2042	ATGAAATGGAAGAGTGTCT 2059			

Db	5458	ATGAGATGGAAGAGTGCT	5475
RESULT 5			
AY615798			
LOCUS	8791 bp	RNA	linear
DEFINITION	Hepatitis C virus isolate HCV.FWB.1	polyprotein gene, partial cds.	
ACCESSION	AY615798		
VERSION	AY615798.1	GI:48479029	
KEYWORDS			
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	1 (bases 1 to 8791)		
AUTHORS	Brann,T.W., Kottlilil,S., Polis,M. and Imamichi,T.		
TITLE	Identification of mutations associated with interferon resistance in HCV and HIV co-infected patients		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 8791)		
AUTHORS	Brann,T.W., Kottlilil,S., Polis,M. and Imamichi,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-MAY-2004) LHR/CSP, SAIC-Frederick, Inc, Building 550, Room 126, 1050 Boyles Street, Frederick, MD 21702, USA		
FEATURES	Location/Qualifiers		
source	1..8791		
	/organism="Hepatitis C virus"		
	/mol_type="genomic RNA"		
	/strain="HCV1a"		
	/isolate="HCV.FWB.1"		
	/db_xref="taxon:11103"		
	/country="USA"		
	67..>8791		
	/codon_start=1		
	/product="polyprotein"		
	/protein_id="AA744836.1"		
	/db_xref="GI:48479030"		
CDS	/translation="MSTNPKPQRKTKRNTNRRPODVKPPGGQIGVGVILLPRRGPRLGVARKTSERSQPRGRPTPKARPEGRTPWAQGYWPLVYNGEGCWAGWLLSPRGSRPSQPTDPRRRSNGKVLDTLTTCGFDLMGYILVAGPLGGAARALAHGVRLVEDGNYATGNLPCGSRFSIFLLALSLCTVPASAYQVRNSGLYHHTNDCPNSSIVYETADAILHSPGCIIPCREGNAKSWAVATPRTDRLGSTQLRRHIDLLVGSATLSALYVGLDCSPFLVQLGFTSPRHHWTQPCNCSILYPGHITGRHMAWDMNMSPTTALVAGLLIPQAILDMTAGAHGVLGAGIYFVSMVGNWAKVLLVLLPAGVDAETSTTGGAAARTVARTSPFAPGPKQNVOLITYNSGSHINRTALNCASLDTGMLAALFYVNSFNSCGERWASCRPLASFDGCGPISYANGSGPEHRYCWHPKPGIYVPAQSGVCGPVYCFTPSVAVGTGKGVTPYNGANDTVFLVNLNTRPFLGNWFGCTWNSTGTFKUGCAPPCVGVGNVTLHCPTDCFRKHPEATYSRGSGFWITPRLVDIYVRLWHVPTCTINYTLFKVRVYGNHLEACNWRGERCDLDRDSLELPLLTITTTQWVLPCSFITLLPALTTGLHLNLIIVDVOYLXGVGSSIVSWAIKWEYVILLFLLADARICSLWMMLLISOAEALNLIVLNNAAGCTHGLASPLVFFCPAWLKGKWPVGMAYAFPGMWPLLLLLALPORAYALDTEMAKSCGVVLVGLMALTISPHYKRVISNCLWLOVFLTRAEALLQVWPLNVRGRDVAILLMCVHVHPLAFDITKLLAALGPLWILQTSLLKVPYFVRVQGLIRICALARKWAGHYVMAIIKAGALTGYIYNHLPLRDWAHNGLDRLDAVAVEPVFQSMETKLTITWAGDPAACGDIINGLIPVSARRGREILLIPADGMSKGRWLLAPITAYAQTRGLLGCITITSLTRGDKNOVEGIIQIVSTAAQTFLATCINGVCTVTVHGAQTRTIASPKGVPMYTNIDOLVGPAPOGARSLTPTCTGSSDLVLTVRHADVI PVRRRGDSRGLSLSPRI SVLYLKGSGGAPLLCPAGHANGI PRAACVTRGVAKADVP I PVNLETTNRSVPTDNSSGPVAPVQSFVAHLHAPTGSKSTKVPAAAYAAQYKVLVNLNPSVAATLFGFAYMSKAHGVDPNIRTVGRTITTPGSPITVSTYKFLADGGCSGAYDIIICDECHSDATSIILIGITVLDQOATAGARLVVLATAPPGSI TVPHENI BEVALSTTGEIPFYGKAI PLBAIKGGRHSLIPCHSKKCDDELAALVALGINAVAYVRLDVSIVPTSGDVVVGA TALMTGTGDRSDIIVDNCVTQDVFSLDPTFTIETTTLPDQAVSRTOQRGRTGGRGKPGIYRVAPGERSGHFDSSVLCEYDAGCAWVELTPAETTVLRAYNTPGLPVQQLDHLFEWFGTGLTHIDAHFUSQTKSGENLPYLVAQVATCARAQPMDQMMKCLIRLPLTHGPTLRLGAVQNVETLTHPVTKYIMTCSADLEVSTVWLVGVYLAALAAVCLSTGVIIIRGLSKGPAVIPDRELYRBEDEMECSOHLPIYEQAGMLIAE QOKALGLLOTASRQAEVIAPAVQTNQWLEAFWAKMWNFI SGIQYLAGLSTLPGNPAISLMAFTAAVSPLTSTOTLLFNILGWSVAQLAAGCAATAFYVAGLAGAAGISVGLGKVLVDLAGYAGVAGALVA FKIMSGRPTTDLVNLPAILSPGALVGVVCAALILRRHVGPBGAVQMMNRLIAFASRGHNSPTHYVPSDDAARVATILSLVTLXLLRLHWOVSECTTPCSGSWIRIDWMI CEVLSDFWTLKAKMLPQLGIPFVSGQRGRI GVMRGDGIWHTRHCHGAEITGHVKNWTRIVGPKTCRNMSGTFFPINAYTGPCTPLP		

QY 181 GGAACAAGGACCATTTGGCTGACCTTAAGGGTCTCTGTTATPCAGATGTACACCAATGTGGAC 240
Db 2170 GGAACGAGGACCATTCGGGTGACCCCAAGGGTCTCTGTCATCCAGATGTATACCAATGTAGAC 2229
QY 241 CAAGACCTCGTAGGCTGGCCGGCTCCCAAGGGTCCCGCTCATTAACACCATGTCACATTGC 300
Db 2230 CAAGACCTTTGGGGCTGGCCGGCTTCGCAAGGTACCCGCTCAITTGACACCCCTGCACATTGC 2289
QY 301 GGCTCTCCGACCTTTTACCTGGTCAAGGACGACCCGATGTCAATCTCTGTGCGCGGACGG 360
Db 2290 GGCTCTCCGACCTTTTACCTGGTCAAGGACGACCCGATGTCAATCTCTGTGCGCGGACGG 2349
QY 361 GGTGATGGAGGGGACGCTCTTTGCGCCCGGGCTATCTCTTACTTGAAGGCTCTCTCG 420
Db 2350 GGTGATAGAGGGGACGCTCTCTGCGCCCGGGCCATTTCTTACTTTGAAGGCTCTCTCG 2409
QY 421 GGAGGCCCTCGCTGTCGCCCGGAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGC 480
Db 2410 GGGGGTCCGCTGTTGTGCCCGCGGGGACGCCGTGGGCATATTTAGGGCCGCGGTGTC 2469
QY 481 ACCCGTGGAGTGGCTAAGGGGCTGGACTTTCATCCCGCTAGAGAGCTTAGAGACAACCATG 540
Db 2470 ACCCGTGGAGTGGCTAAGGGGCTGGACTTTCATCCCTGTGGAGAACCTTAGAGACAACCATG 2529
QY 541 AGTCCCGGGTGTCTCAGACAACTCTCCCAACAGAGTGGCCCGGAGCTACCAAGTG 600
Db 2530 AGTCCCGGGTGTCTCAGAGTAACTCTCTCCACAGTAGTGGCCCGGAGCTTCCAGGTG 2589
QY 601 GCCCACCTGCATGCTCCACCGGAGGCTAAGGACCAAGGTCCCGCGCGCATACGCA 660
Db 2590 GCTCACCTGCATGCTCCACAGGAGGCGGCAAAAGCAACCAAGGTCCCGCGCTGCATATGCA 2649
QY 661 GCTCAGGGCTACAAGGTGCTGGTCTCAACCCCTCCGTGTGTCACAAATGGGCTTTGGT 720
Db 2650 GCTCAGGGCTACAAGGTGCTGGTCTCAACCCCTCCGTGTGTCACAAATGGGCTTTGGT 2709
QY 721 GCTTACATGTCGAAGGGCCATGGGATGATCTTAACATCAGGACTGGGGTGAGGACAATT 780
Db 2710 GCTTACATGTCGAAGGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 2769
QY 781 ACTACTGGCAGCCGATCATCGTATTCACCTACGGCAAGTTCCTTGGCGAGCGGGGTG 840
Db 2770 ACCACTGGCAGCCGATCATCGTATTCACCTACGGCAAGTTCCTTGGCGAGCGGGGTG 2829
QY 841 TCAGGGGTGCTATGACATAAATTTGTGACAGTGCACCTCCACGATGCAACATCC 900
Db 2830 TCAGGGGTGCTATGACATAAATTTGTGACAGTGCACCTCCACGATGCAACATCC 2889
QY 901 ATCTTGGGCAATGGCACTGTCTTGACCAAGCAGAGACCGGGGGGAGACTGACATGTG 960
Db 2890 ATCTTGGGCAATGGCACTGTCTTGACCAAGCAGAGACTGCGGGGGGAGACTGCTTTGTG 2949
QY 961 CTGCGCAACGGTACCCCTCGGGTCTGTCATGTCGCCCATCTTAACATCGAGGAGGTT 1020
Db 2950 CTGCGCAACGGTACCCCTCGGGTCTGTCATGTCGCCCATCTTAACATCGAGGAGGTT 3009
QY 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGCGAAGGCTATTTCCCTTGAAGCAATT 1080
Db 3010 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGCGAAGGCTATTTCCCTTGAAGTATC 3069
QY 1081 AAGGGGGGAGACATCTCATCTTCTGCACTCAAGGAAGAGTGCGAGCTGCGCGCA 1140
Db 3070 AAGGGGGGAGACATCTCATCTTCTGCACTCAAGGAAGAGTGCGAGCAACTGCGCGCA 3129
QY 1141 AAATGTCGGTGGGTCATATGCGGTGCTTACTACCGCGGCTTGTATGTCGCTC 1200
Db 3130 AAGTGTGTCATTTGGGTCATCAATGCGGTGCTTACTACCGGCTTGTATGTCGCTC 3189
QY 1201 ATCCCGACGAGTGTGAGTGTGTCGTGTCGCACTGACGCCCTCATGACCGGCTTTACC 1260
Db 3190 ATCCCGACGAGTGTGAGTGTGTCGTGTCGCAACCGATGCTTCTCATGACCGGCTATACC 3249

QY 1261 GGCGACTTCGATTCGGGTAGACTGCAACACGTTGTGTCAACCCAGACAGTCGACTTCAGC 1320
Db 3250 GGCGACTTCGACTCGGTAGACTGCAATACGTTGTGTCAACCCAGACAGTCGACTTCAGC 3309
QY 1321 CTTGACCTTACCTTCCATTGAGACAATCAGCTTCCCGAGGATGTCTCTCCGCTACT 1380
Db 3310 CTTGACCTTACCTTCCATTGAGACAATCAGCTTCCCGAGGATGTCTCTCCGCTACT 3369
QY 1381 CAACGTCCGGGTAGGACTGGCAGAGGAAGCCAGGCACTTACAGATTGTGGCAACCGGG 1440
Db 3370 CAACGTCCGGGGAGGACTGGCAGGGGAAGCCAGGCATCTACAGATTGTGGCAACCGGG 3429
QY 1441 GAGGCTCTTCTGGCATGTTTGTACTGTCGTCTCTCTGCGAGTGTATGACCGCGGTTGT 1500
Db 3430 GAGCGCCCTTCCGGCATGTTTGTACTGTCGTCTCTCTGAGTGTATGACGCGAGGCTGT 3489
QY 1501 GCTTGTGATGAGCTTACGCCCGGAGACCAAGTGTAGGCTTACGAGCATACATGAACACC 1560
Db 3490 GCTTGTGATGAGCTTACGCCCGGAGACTTACAGTTAGGCTACGAGGCTACATGAACACC 3549
QY 1561 CCGGACTTCCCGGTGTCGCAAGACCATCTTTGAATTTTGGAGGGCGCTCTTTACGGGTCTC 1620
Db 3550 CCGGACTTCCCGGTGTCGCAAGACCATCTTTGAATTTTGGAGGGCGCTCTTTACAGGCTC 3609
QY 1621 ACCCAGATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAAAACCTTCCCTAT 1680
Db 3610 ACTCATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTCTTAC 3669
QY 1681 CTGCTAGCTACCAAGCCCGGTGTCGCTAGAGCTCAAGCCCTCCCGTCTCGTGGGAC 1740
Db 3670 CTGCTAGCTACCAAGCCCGGTGTCGCTAGAGCTCAAGCCCTCCCGTCTCGTGGGAC 3729
QY 1741 CAGATGTGGAAGTCTTGTATCCGCTCTCAAGCCCAACCTCCATGGGCCCAACACCTCTGCTA 1800
Db 3730 CAGATGTGGAAGTCTTGTATCCGCTCAAGCCCAACCTCCATGGGCCCAACACCTCTGCTA 3789
QY 1801 TATAGACTGGGGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAAATATATC 1860
Db 3790 TACAGACTGGGGCTGTCTCAGAAATGAATCAACCTGACGACCCAGTCAACAAATATATC 3849
QY 1861 ATGACATGATGTGCGGTGACCTGAGTGTGTCGCTGTCAGAGTACCTGGGTCTGTTGGCGGC 1920
Db 3850 ATGACATGATGTGCGGTGACCTGAGTGTGTCGCTGTCAGAGTACCTGGGTCTGTTGGCGGC 3909
QY 1921 GTTCTGCTCTTGTGGCGGTATTTGCTATCCACAGGCTGCTGCTCATAGTAGGTAGG 1980
Db 3910 GTTCTGCTCTTGTGGCGGTATTTGCTATCCACAGGCTGCTGCTCATAGTAGGTAGG 3969
QY 1981 ATTGTCTTGTCCGGAAGCCGCAATCATACCCGACAGGGAAGTCTCTTACCGGGAGTTC 2040
Db 3970 GTGCTCTTGTCCGGAAGCCGCAATCATACCTGACAGGGAAGTCTCTTACCGGAGTTC 4029
QY 2041 GATCAATGCAAGTGTCT 2059
Db 4030 GATGAGATGGAAGTGTCT 4048

RESULT 7
I06434
LOCUS I06434
DEFINITION Sequence 48 from Patent EP 0318216.
ACCESSION I06434
VERSION I06434.1 GI:590311
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 5360)
AUTHORS Houghton, M., Choo, Q.-L., and Kuo, G.
TITLE Nanbv diagnostics and vaccines
JOURNAL Patent: EP 0318216-A1 48 31-MAY-1989;
FEATURES Location/Qualifiers
source 1..5360

ORIGIN		/organism="unknown" /mol_type="unassigned DNA"		Query Match Best Local Similarity 86.7%; Score 1786; DB 6; Length 5360; Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;			
QY	2	TGGCGCCTATACAGGCTATGCCAGCAGACAAAGGGCCCTTTGGGATGCATAATCACCA	61	Db	1891	TCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGGAGTTG	1950
Db	931	TGGCGCCCATACGGCGTACGCCAGCAGACAAAGGGCCCTCTAGGTGCTAATCACCA	990	QY	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCTCTTGAACAATTA	1081
QY	62	GCTTTGACCGGCGGGACAAAACACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG	121	Db	1951	CTCTGTGCCACCAACCGGAGAGATCCCTTTTATCGGCAAGGCTATCCCTCTCGAAGTAATCA	2010
Db	991	GCCTAACTGGCGGGACAAAACCAAGTGGAGGTGAGGTCAGATTGTGTCAACTGCTG	1050	QY	1082	AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAAAGTGCAGAGCTCGCCGCAA	1141
QY	122	CCGAGACTTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCGG	181	Db	2011	AGGGGGGAGACATCTCATCTTCTGTCTATTCAAAGAAAGTGCAGAGCTCGCCGCAA	2070
Db	1051	CCCAACCTTCTTGGCAACGTGTCAATCAATGGGGTGTGTGGACTGTCTACCAACGGGGCGG	1110	QY	1142	AACGTGTCCGTGGCGTCAATGCGTGGCTTACTACCGCGGCTTGTATGTGTCCGTCA	1201
QY	182	GAACAAGGACCATTTGGGTCACTTAAGGGTCCCTGTTATCCAGATGTACACCAATGTGACC	241	Db	2071	AGCTGTCCGATTTGGGCATCAATGCGTGGCTACTACTACCGCGTCTTGAACGTGTCCGTCA	2130
Db	1111	GAACGAGGACCATCGCGTCAACCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGACC	1170	QY	1202	TCCGACCAAGTGTGACGCTGTGTGTGTGGCAACTGAACGCCCTCATGACCGGCTTTACCG	1261
QY	242	AAGACCTGTAGCTGGCGCGCTCCCAAGGTGCCCGTCAATTAACAACATGCACTTGGC	301	Db	2131	TCCGACCAAGCGGCGATGTGTGTGTGTGGCAACCGATGCCCTCATGACCGGCTATACCG	2190
Db	1171	AAGACCTTGTGGCTGGCGCGCTCCCAAGGTAGCGGCTCAATTGACACCTTGCACTTGGC	1230	QY	1262	GGGACTTCGATTCGGGTGATAGACTGCAACACGTGTGTCAACCCAGACAGTTCAGACC	1321
QY	302	GCTCTCGGACCTTTACTGTGTACGAGGACGCGGATGTCATTCCTGTGCGCGGAGCGG	361	Db	2191	GGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCAACCCAGACAGTTCGATTTACGCC	2250
Db	1231	GCTCTCGGACCTTTACTGTGTACGAGGACGCGGATGTCATTCCTGTGCGCGGAGCGG	1290	QY	1322	TTGACCTTACCTTACCATTTGAGACAAATCAGCTTCCCAGGATGCTGTCTCCCGTACTC	1381
QY	362	GTGATGACGGGCGAGCTGCTTTTCCGCCCGGCTATCTTACTTGAAGGCTCTCTCGG	421	Db	2251	TTGACCTTACCTTACCATTTGAGACAAATCAGCTTCCCAGGATGCTGTCTCCCGCACTC	2310
Db	1291	GTGATGACGGGCGAGCTGCTGCTCGCCCCGGCCCAATTTCTTACTTTGAAGGCTCTCTCGG	1350	QY	1382	AACGTCTGGGTAGGACTGGCAGAGGGAAGCCAGGCAATCTACAGATTTTGTGGCACCGGGG	1441
QY	422	GAGGCCCTCTGTGTGCCCCCGCAGGACATCCCGTAGGCATATTCAGAGCCCGGTATGCA	481	Db	2311	AACGTCTGGGCGAGGACTGGCAGGGGGAAGCCAGGCAATCTACAGATTTTGTGGCACCGGGG	2370
Db	1351	GGGTCTCGCTGTTGTGCCCCCGGGGACGCGGTGGGCATATTTAGGGCCCGGTGTGCA	1410	QY	1442	AGCGTCTCTTCTGGCATGTTTGACTCGTCTGCTCCTCTCGAGTGTCTATGACGCGGTTGTG	1501
QY	482	CCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGCTTAGAGCAACCATGA	541	Db	2371	AGCGGCCCTCTCGGCATGTTTCCACTCGTCTCGTCTCTGTGAGTGTCTATGACGCGGCTGTG	2430
Db	1411	CCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCTGTGGAGAACCTAGAGCAACCATGA	1470	QY	1502	CTTGTGTATGAGCTTACGCCCGCGAGACACAGTTAGGCTACGAGCATACATGAACACCC	1561
QY	542	GGTCCCCGGTGTCTCAGACAACTCTCTCCCAACAGCAGTGCAGAGTCCCAAGTGG	601	Db	2431	CTTGTGTATGAGCTACGCCCGCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACCC	2490
Db	1471	GGTCCCCGGTGTCTCAGGATTAECTCTCTCCACAGTAGTAGTGCACAGCTTCCAGGTGG	1530	QY	1562	CGGACTTCCCGTGTGCGAAGCAACCATCTTGAATTTTGGAGGGGCTCTTTACGGGTCTCA	1621
QY	602	CCCACTGTATGCTCCCAACGGGAGGTTAAGACCAAGGTCCCGGCGCATACGAG	661	Db	2491	CGGGGCTTCCCGTGTGCGCAGGACCATCTTGAAATTTTGGAGGGGCGTCTTTTACAGGGCTCA	2550
Db	1531	CTCACCTCATGCTCCACAGGCGAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCAG	1590	QY	1622	CCCATAGAGCCGCACCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTTCCCTATC	1681
QY	662	CTCAGGGCTACAGGTGTGCTGCTCAACCCCTCCGTGCTGCTGCAACATGGGCTTTGGTG	721	Db	2551	CTCATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCTCTTACC	2610
Db	1591	CTCAGGGCTATAGGTGTGTACTCAACCCCTCTGTGTGCTGCAACACTGGGGCTTTGGTG	1650	QY	1682	TGGTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCCCTCCCGCTCGTGGGACC	1741
QY	722	CTTACATGTCCAAAGGCCATGGGATTTGATCCTTAACATCAGGACTGGGTGAGGACAATTA	781	Db	2611	TGGTAGCGTACCAAGCCACCGTGTGCGCTAGGCTCAAGCCCCCTCCCGCTCGTGGGACC	2670
Db	1651	CTTACATGTCCAAAGGCTCATGGGATGCTATCTTAACATCAGGACCGGGGTGAGAACAAATTA	1710	QY	1742	AGATGTGAAGTGTGATCGCTCTCAAGCCCACTTCAAGGGCCCAACCTCTGCTGCTAT	1801
QY	782	CTACTGGAGCCCGATGACGTATTCACCTACGGCAAGTTCTCTTTCGCCAGCGCGGGTGT	841	Db	2671	AGATGTGAAGTGTGATTCGCTCAAGGCCCACTTCCATAGGGCCCAACACCTCTGCTAT	2730
Db	1711	CCACTGGAGCCCATACGTACTCCACTACGGCAAGTTCTTTCGCCAGCGGGGTGCT	1770	QY	1802	ATAGACTCGGCGCTGTCAGAAATGAAAGTCAAGTCAACCTGACGCCAGTCCCAAGTATATCA	1861
QY	842	CAGGGGTGCTTATGACATAATTTGTGACGAGTGCCACTCCACGGATGCAACATCCA	901	Db	2731	ACAGACTGGGCGCTGTTTCAAGATGAATCACTCTGACGCAACCGTCTACCAATATATCA	2790
Db	1771	CGGGGGCGCTTATGACATAATTTTGTGACGAGTGCCTTCCACGGATGCCATCCA	1830	QY	1862	TGACATGTATGTCCGCTGACCTGGAGGTGCTCACGAGTACCTGGGTGCTCGTTGGCGGG	1921
QY	902	TCTTGGGCATTTGCACTGTTCCTTACCAAGCAGAGACCGGGGGGAGACTGACTGTGC	961	Db	2791	TGACATGTATGTCCGCGCATCTGGAGGTGCTCACGAGCACCTGGGTGCTCGTTGGCGGG	2850
Db	1831	TCTTGGGCATTCGCACTGTCTTGACCAAGCAGAGACTGTCCGGGGGCGAGACTGGTTGTC	1890	QY	1922	TTCTGGCTGCTTTTGGCGCGGTATTTGCCATTTCCACAGGCTCGTGTGTCTATGTAGTAGGA	1981
QY	962	TGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCTCTTAACATCGAGGAGTTG	1021	Db	2851	TCCTGGCTGCTTTTGGCGCGGTATTTGCCGTGTCAACAGGCTGCTGTGTGTAGTGGGCGGG	2910
QY				QY	1982	TTGTCTTGTCCGGAAGCCGCAATATATCCGACAGGGAAGTCTCTCTACCGGGAGTTTCG	2041
				Db	2911	TCGTCTTGTCCGGGAAGCCCGCAATCATCTGACAGGGAAGTCTCTTACCGGAGAGTTTCG	2970
				QY	2042	ATGAAATGGAAGAGTGTCT 2059	

Db 2971 ATGAGTGGAGAGTGCT 2988

RESULT 8
109328
LOCUS 109328 5360 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 8 from Patent WO 8904669.
ACCESSION 109328
VERSION 109328.1 GI:587963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5360)
AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 8 01-JUN-1989;
FEATURES
source 1..5360
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 86.7%; Score 1786; DB 6; Length 5360;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGCGCCTATCACGGCCTATGCCAGCAGACAAGGGGCGCTTTTGGGATGCATATACCA 61
Db 931 TGGCGCCCATCACGGCTACGCCCAGCAGACAAGGGGCGCTTAGGGTGCATATACCA 990

Qy 62 GCTTGACGGCGGGACAAMACCAGGTGAGGGTGAAGTTGAGATCGTGTCAACTGCTG 121
Db 991 GCCTAACTGGCGGGACAAMACCAGGTGAGGGTGAAGTTGAGATCGTGTCAACTGCTG 1050

Qy 122 CCGAGACTTTCTTGGCAACCTGCAATTAACGGGTGCTGTGGACTGTCTACCATGGAGCG 181
Db 1051 CCGAATCTTCTGGCAACCTGCAATTAACGGGTGCTGTGGACTGTCTACCATGGAGCG 1110

Qy 182 GAAACAAGGACCATTCGCTACCTAAGGGTCTCTGTATCCAGATGTACACCAATGTGACC 241
Db 1111 GAACAGGACCATTCGCTACCTAAGGGTCTCTGTATCCAGATGTATACCAATGTGACC 1170

Qy 242 AAGACTCTGATGGTGGCGGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGGC 301
Db 1171 AAGACTCTGATGGTGGCGGCTCCCAAGGTGGCGCTCATTAACACCATGCACTTGGC 1230

Qy 302 GCTCCTCGGACCTTACCTGTGTACAGGACGCGGATGCTATTCCTGTGGCGGCGG 361
Db 1231 GCTCCTCGGACCTTACCTGTGTACAGGACGCGGATGCTATTCCTGTGGCGGCGG 1290

Qy 362 GTGATGCGAGGGCAGGCTGTCTTTCGCGCGGCGCTATCTTACTTGAAGGCTCTCTCG 421
Db 1291 GTGATGCGAGGGCAGGCTGTCTTTCGCGCGGCGCTATCTTACTTGAAGGCTCTCTCG 1350

Qy 422 GAGGCCCTCTGTGTGCCCGCAGGAATGCGGTAGGCAATTCAGAGCGCGGTATGCA 481
Db 1351 GGGTCCGCTGTGTGCCCGCAGGAATGCGGTAGGCAATTCAGAGCGCGGTATGCA 1410

Qy 482 CCGGTGAGTGGCTTAAGCGGTGACTTATCCCGGTAGAGAGCTTAGACACACCATGA 541
Db 1411 CCGGTGAGTGGCTTAAGCGGTGACTTATCCCGGTAGAGAGCTTAGACACACCATGA 1470

Qy 542 GGTCCCGGTGTCTCAGACAACCTCCCTCCACAGAGTGCCTCCAGAGCTTACCAAGTGG 601
Db 1471 GGTCCCGGTGTCTCAGACAACCTCCCTCCACAGAGTGCCTCCAGAGCTTACCAAGTGG 1530

Qy 602 CCGACCTGCACTGTCCACCGGACGCGGTAAAGAGCAACCAAGGTCCCGCGGCTATGCGAG 661
Db 1531 CTACCTCCATGTCTCCACAGGACGCGGCAAAAGCAACCAAGGTCCCGGCTGCAATGCA 1590

Qy 662 CTAGGCGTCAAGGTGCTGTGTCAACCGCTCCGTTGCTGCAAAATGGGCTTTGGTG 721
Db 1591 CTCAGGCGTCAAGGTGCTGTGTGTCAACCGCTCCGTTGCTGCAAAATGGGCTTTGGTG 1650

Qy 722 CTTACATGTCCAAGGCCATGGGATGGATCCTTAACATCAGGACTGGGGTGAAGCAATTA 781
Db 1651 CTTACATGTCCAAGGCTCATGGGATGGATCCTTAACATCAGGACCGGGGTGAGAACAATTA 1710

Qy 782 CTACTGGCAGCCGATCAGCTATTCACCTACGGCAAGTTCCTTGGCGGCGGGGTGTT 841
Db 1711 CCACTGGCAGCCGATCAGCTATTCACCTACGGCAAGTTCCTTGGCGGCGGGGTGCT 1770

Qy 842 CAGGGGTGCTTATGACATAAATTTGTGACAGTGCCTTCCACCGATGCAACATCCA 901
Db 1771 CGGGGGCGCTTATGACATAAATTTGTGACAGTGCCTTCCACCGATGCAACATCCA 1830

Qy 902 TCTTGGCATTTGGCACTCTCTTGAACAAGCAGAGACCGGGGGCGAGACTGACTGTC 961
Db 1831 TCTTGGCATTCGGCACTCTCTTGAACAAGCAGAGACTGCGGGGGCGAGACTGCTGTC 1890

Qy 962 TCGCCACCGCTACCCCTCCGGGCTCCGTCAGTGTGCCCATCTTAACATCAGGAGGTG 1021
Db 1891 TCGCCACCGCCACCCCTCCGGGCTCCGTCAGTGTGCCCATCTTAACATCAGGAGGTG 1950

Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTTGAACAATTA 1081
Db 1951 CTCTGTCCACACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTTGAAGTATCA 2010

Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTGCAGAGCTCCGCCAA 1141
Db 2011 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTGCAGAGCTCCGCCAA 2070

Qy 1142 AACTGTGCGGTTGGGGTCAATGCCGTGCTTACTACCGGGCTTGTGTGCTGCTCA 1201
Db 2071 AGCTGTGCTATTTGGGCTCAATGCCGTGCTTACTACCGGGCTTGTGTGCTGCTCA 2130

Qy 1202 TCCGACACAGTGTGCTGTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1261
Db 2131 TCCGACACAGTGTGCTGTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2190

Qy 1262 GCGACTTCGATTCGGTGATAGACTGCAACACAGTGTGTCACCCAGACAGTGCAGCTTACGCC 1321
Db 2191 GCGACTTCGATTCGGTGATAGACTGCAACACAGTGTGTCACCCAGACAGTGCAGCTTACGCC 2250

Qy 1322 TTGACCTTACCTTACCATTTGAGACATCAGCTTCCCGAGGATGCTGTCTCCGCTACTC 1381
Db 2251 TTGACCTTACCTTACCATTTGAGACATCAGCTTCCCGAGGATGCTGTCTCCGCTACTC 2310

Qy 1382 AAGTGGGGTAGACTTGGCAGAGGGAAGCAGGCACTTACAGATTTGTGCGACCGGGGG 1441
Db 2311 AAGTGGGGTAGACTTGGCAGAGGGAAGCAGGCACTTACAGATTTGTGCGACCGGGGG 2370

Qy 1442 AGCGTCCCTTCTGGCATGTTTGAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
Db 2371 AGCGCCCTTCCGCGATGTTTCGACTCGTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2430

Qy 1502 CTTGTGATGAGCTTAGCCCGCGGAGACACAGTGTAGGCTTACAGAGCATATAGTAACACC 1561
Db 2431 CTTGTGATGAGCTTAGCCCGCGGAGACACAGTGTAGGCTTACAGAGCATATAGTAACACC 2490

Qy 1562 CGGGAATTCCTGTCGCAAGACATCTTGAATTTTGGGAGGCGCTTTTACGGGCTCTCA 1621
Db 2491 CGGGGCTTCCCGTGTGCGAGGACATCTTGAATTTTGGGAGGCGCTTTTACAGGCTCTCA 2550

Qy 1622 CCACATAGACCGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAAGAACTTCCCTTATC 1681
Db 2551 CTATATAGTATGCCACTTCTATCCAGACAAAGCAGAGTGGGGAAGAACTTCCCTTATC 2610

Qy 1682 TGTAGCTACCAAGCACAAGTGTGCTGCTAGAGCTCAAGCCCTTCCCTCCCTGCTGGGACC 1741
Db 2611 TGTAGCTACCAAGCACAAGTGTGCTGCTAGAGCTCAAGCCCTTCCCTCCCTGCTGGGACC 2670

Qy 1742 AGATGTGAGTCTTGTATCCGTCCTCAAGCCACCTCCATGGCGGCAACACTCTGCTAT 1801
Db 2671 AGATGTGAGTCTTGTATCCGTCCTCAAGCCACCTCCATGGCGGCAACACTCTGCTAT 2730

Qy 1802 ATAGACTGGCGCTGTCAGAGTAAGTCAACCTGACGACCCAGTACCAAGTATATCA 1861
Db 2731 ACAGACTGGCGCTGTTTCAAGTGAATCAACCTGACGACCCAGTACCAAAATACATCA 2790
Qy 1862 TGACATGATGTCGGCTGACCTGGAGTCTGACAGTACCTGGGTGCTGTTGGCGCG 1921
Db 2791 TGACATGATGTCGGCGACCTGGAGTCTGACAGTCTGCTGGTGTCTGTTGGCGCG 2850
Qy 1922 TTCTGGCTGCTTTGGCGCGCTATTGCTTATCCACAGGCTCGTGGTTCATAGTAGTAGGA 1981
Db 2851 TCCTGGCTGCTTTGGCGCGCTATTGCTTCAACAGGCTCGTGGTTCATAGTAGTAGGA 2910
Qy 1982 TTGTCTTGTCCGAAAGCCCGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTCG 2041
Db 2911 TCGTCTTGTCCGGAAGCCCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTCG 2970
Qy 2042 ATGAATGGAAGATGCT 2059
Db 2971 ATGAGATGGAAGATGCT 2988

RESULT 9
I06440
LOCUS I06440 6785 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 54 from Patent EP 0318216.
ACCESSION I06440
VERSION I06440.1 GI:590312
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6785)
AUTHORS Houghton,M., Choo,Q.-L. and Kuo,G.
TITLE Nanbv diagnostics and vaccines
JOURNAL Patent: EP 0318216-A1 54 31-MAY-1989;
FEATURES
source
1..6785
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 86.7%; Score 1786; DB 6; Length 6785;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
Qy 2 TGGCGCTATACAGGCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCAATACCA 61
Db 1204 TGGCGCCATACAGGCTATGCCAGCAGACAAAGGGGCTCTTAGGGTGCAATACCA 1263
Qy 62 GCTTGACCGCGCGGACAAACACAGGTGGAGGTGAGGTTCAGATCGTGTCAACTGCTG 121
Db 1264 GCCTAACTGGCGCGGACAAACCAAGTGGAGGTGAGGTTCAGATGCTCAACTGCTG 1323
Qy 122 CCAGACTTTCTTGGCAACTGATTAACGGGTGTGTTGGACTGTCTACCATGGAGCGG 181
Db 1324 CCCAAACCTTCTGGCAACTGATCAATGGGGTGTCTGGAAGTGTCTACCGCGCGCG 1383
Qy 182 GACAGGACCATTTGGCTGACCTAAGGTCCTGTTATCCAGATGTACCAATGTGGACC 241
Db 1384 GAACGAGGACCATTCGGCTGACCAAGGTCCTGATCCAGATGTATCAATGTAGACC 1443
Qy 242 AAGACCTCGTAGGCTGGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGGC 301
Db 1444 AAGACCTTGTGGCTGGCGCTCCGCAAGGTAGCGCTCATTTGACACCTGCACTTGG 1503
Qy 302 GCTCTCGGACCTTTACCTGGTCAAGGACCGCCGATGATCTCTGTGGCGCGACGGG 361
Db 1504 GCTCTCGGACCTTTACCTGGTCAAGGACCGCCGATGATCTCTGTGGCGCGCGG 1563
Qy 362 GTGATGGCAGGGCAGCTGCTTTCGCCCGGCTATCTCTTACTTGAAGGCTCTCTCGG 421
Db 1564 GTGATGACGGGCGAGCTGCTGTCGCCCGGCGCAATTTCTTACTTTGAAGGCTCTCTCGG 1623

Qy 422 GAGGCGCTCTGTGTGTCGCCCGCAGGACATGCGGTAGGCATATTTAGAGCCGCGGTATGCA 481
Db 1624 GGGGTCCGCTGTGTGTCGCCCGCGGGGCAAGCGGTGCGCATATTTAGGGCCGCGGTGCA 1683
Qy 482 CCGGTGGAGTGGCTAAGGGCGGTGGACTTCATCCCGGTAGAGACTTTAGAGACAAACATGA 541
Db 1684 CCGGTGGAGTGGCTAAGGGCGGTGGACTTCATCCCGGTGGAGAACCTTAGAGACAAACATGA 1743
Qy 542 GGTCCCGCGGTGTTCTCAGACAACTCCTCCCAACAGCAGTGGCCCAAGACTACCAAGTGG 601
Db 1744 GGTCCCGCGGTGTTACGGATAACTCTCTCCACCAAGTAGTGGCCCAAGACTTCCAGTGG 1803
Qy 602 CCCACCTGCATGCTCCCAACCGCAGCGGTAAAGACCAAGGTTCGCCGCGCGCATAGCGAG 661
Db 1804 CTCACCTCCATGCTCCCAAGCAGCGGCAAAAGCACAAGGTTCGCCGCTGCATATGCGAG 1863
Qy 662 CTCAGGGCTACAAGGTGCTGGTCTCAACCCCTCCGTTGCTGCTGCAACAATGGGGCTTGGTG 721
Db 1864 CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTGGTG 1923
Qy 722 CTTACATGTCCTAAGGCCCATGGATTGATCTTAACATCAGGACTGGGGTGAGACAATTA 781
Db 1924 CTTACATGTCCTAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGACAATTA 1983
Qy 782 CTACTGGCAGCCCGATCAGTATTCACCTACGGCAAGTTCTTGGCCGACGGCGGTGT 841
Db 1984 CCACTGGCAGCCCATCAGTACTCCACCTACGGCAAGTTCTTGGCCGACGGCGGTGT 2043
Qy 842 CAGGGGTGCTTATGACATAATAATTGTTGACGAGTGCCACTCCACGGATGCAACATCCA 901
Db 2044 CGGGGGCGCTTATGACATAATAATTGTTGACGAGTGCCACTCCACGGATGCCACATCCA 2103
Qy 902 TCTTGGGCAATTGGCACTGTCTTGAACAGCAGAGACCGCGGGCGGACGACTGTGTCG 961
Db 2104 TCTTGGGCAATCGCACTGTCTTGAACAGCAGAGACTGCGGGGGGAGACTGTGTTGTC 2163
Qy 962 TCGCCACGCTACCCCTCGGGCTCCGTCACTGTGCCCATCTTAACATCGAGGAGTTG 1021
Db 2164 TCGCCACGCGCACCCCTCGGGCTCCGTCACTGTGCCCATCTCCCAACATCGAGGAGTTG 2223
Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATTA 1081
Db 2224 CTCTGTCCACACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGTAATCA 2283
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCGACGAGTCTCGCGCAA 1141
Db 2284 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCGACGAGTCTCGCGCAA 2343
Qy 1142 AACTGTCGCGTTGGCGGTCAATGCGGTGGCTTACTACGCGGCGCTTGATGTCCGTCA 1201
Db 2344 AGCTGTCGCAATTGGGCATCAATGCGGTGGCTTACTACGCGGTCTTGAAGTCCGTCA 2403
Qy 1202 TCCGACACGTCGTGACGTTGTCGTGGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261
Db 2404 TCCGACACGTCGTGACGTTGTCGTGGGCAACGATGCCCTCATGACCGGCTTACCG 2463
Qy 1262 GGGACTTGGATTCGGGTGATAGACTCAACACGCTGTGTCAACCCAGACAGTCTCAGCC 1321
Db 2464 GGGACTTGGATTCGGGTGATAGACTCAACACGCTGTGTCAACCCAGACAGTCTCAGCC 2523
Qy 1322 TTGACCTACCTTCACTTACCAATTGAGACAATCAGCTTTCCCAAGGATGCTGTCTCCCGTACTC 1381
Db 2524 TTGACCTTACCTTCACTTACCAATTGAGACAATCAGCTTTCCCAAGGATGCTGTCTCCCGTACTC 2583
Qy 1382 AACGTGGGGTAGGACTGGCAGAGGGAAGCCAGGATCTACAGATTTGTGGACACCGGGG 1441
Db 2584 AACGTGGGGTAGGACTGGCAGAGGGAAGCCAGGATCTTACAGATTTGTGGACACCGGGG 2643
Qy 1442 AGCGTCTTCTTGGCATGTTTGGATGCTGTCTCTGCGAGTGCTATACGCGGGTGTG 1501
Db 2644 AGCGTCTTCTTGGCATGTTTGGATGCTGTCTCTGCGAGTGCTATGACGAGGCTGTG 2703
Qy 1502 CTTGTGATAGCTTACGCCCGCGGAGACCAAGTTAGGCTACGAGATATACGAAACCCC 1561

```
Db 2704 CTTGGTATGAGCTCACCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACCC 2763
Qy 1562 CGGACTTCCCGTGTGCCAAGAACCATCTTGAATTTTGGAGGGCGTCTTTACGGGTCTCA 1621
Db 2764 CGGGCTTCCCGTGTGCCAAGAACCATCTTGAATTTTGGAGGGCGTCTTTACAGGCTCA 2823
Qy 1622 CCACATAGAGCCCACTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
Db 2824 CTATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACTTCTTACC 2883
Qy 1682 TGGTAGGTCACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
Db 2884 TGGTAGGTCACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 2943
Qy 1742 AGATGTGGAAGTCTTGATCCGCTCAAGCCACCCCTCCATGGGCAACACCTCTGCTAT 1801
Db 2944 AGATGTGGAAGTGTGATGCTTGCCTCAAGCCACCCCTCCATGGGCAACACCCCTGCTAT 3003
Qy 1802 ATAGACTGGGGCTGTCCAGAAATGAATCAACCTTGACGCCACCCAGTCAACCAAGTATATCA 1861
Db 3004 ACAGACTGGGCGCTGTTAGAAATGAATCAACCTTGACGCCACCCAGTCAACCAATATATCA 3063
Qy 1862 TGACATGTATGCTGGCTGACCTGAGGTCTGTCAGAGTACCTGGGTCTGTTGGCGGCG 1921
Db 3064 TGACATGCATGTGCGGCCGACCTGGAGTCTGTCACAGACACCTGGGTCTGTTGGCGGCG 3123
Qy 1922 TTCTGGCTGCTTTGGCGCGGTATTGCTTATCCAGAGCTGCTGCTATAGTAGTAGGA 1981
Db 3124 TCCTGGCTGCTTTGGCGCGGTATTGCTTATCCAGAGCTGCTGCTATAGTAGTAGGAGG 3183
Qy 1982 TTGCTTTGTCGGAAAGCCGGCAATATACCCGACAGGGAAGTCTCTACCGGAGTTG 2041
Db 3184 TCGTCTTGTCCGGAAGCCGGCAATATACCTGACAGGGAAGTCTCTACCGGAGTTG 3243
Qy 2042 ATGAAATGGAAGTGTCT 2059
Db 3244 ATGAGATGGAAGTGTCT 3261

RESULT 10
ARL18696
LOCUS ARL18696 7310 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 74 from patent US 6150087.
ACCESSION ARL18696
VERSION ARL18696.1 GI:14100606
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7310)
AUTHORS Chien,D.Y.
TITLE NANBV diagnostics and vaccines
JOURNAL Patent: US 6150087-A 74 21-NOV-2000;
FEATURES
    Location/Qualifiers
        1..7310
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN

Query Match 86.7%; Score 1786; DB 6; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGGCTATACGGCTATGCCAGCAGACAGAGGGGCTTTGGGATGATATACCA 61
Db 1729 TGGGCCCATACAGGCTATGCCAGCAGACAGAGGGGCTTCTAGGGTGCATATACCA 1788
Qy 62 GCTTGACCGCGCGACAAACCAAGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG 121
Db 1789 GCCTAACTGGCGCGGACAAACCAAGTGGAGGTGAGGTTTCAGATGCTGTAAGTCTG 1848
Qy 122 CCCAGACTTTTGGCAACCTGCATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCG 181
```

```
1849 CCACAACTTCTCCGGCAACGTGATCAATGGGGTGTGCTGGACTGTCTACACGGGCGG 1908
182 GAACAGGACCATTTGGCTCACCTTAAGGTCTCTGTTATCCAGATGTACACCAATGTGACC 241
1909 GAACAGGACCATTCGGCTCACCAAGGTCTCTGATCCAGATGTATACCAATGTAGACC 1968
242 AAGACCTCTGTAAGTGTGCGCCGCTCCCAAGGTGCGGCTCATTAACACCAATGTGCG 301
1969 AAGACCTTGTGGCTGGCGCGCTCCGCAAGGTAGCGCTCATTTGACACACCTGTCACTTGG 2028
302 GCTCTCGGACCTTTACCTGGTCAAGAGGCAAGCGGATGTCTCTGTTGGCGCGGACGG 361
2029 GCTCTCTCGGACCTTTACCTGGTCAAGAGGCAAGCGGATGTCTCTGTTGGCGCGGACGG 2088
362 GTGATGACAGGGGAGCGCTCTGTTGGCGCGGACCGCTCTCTCTGTTGAAAGGCTCTCTCG 421
2089 GTGATGACAGGGGAGCGCTCTGTTGGCGCGGACCGCTCTCTCTGTTGAAAGGCTCTCTCG 2148
422 GAGGCCCTCTGCTGTGCGCCGCGAGGACATCGCTAGGCATATTTACAGAGCGCGGTATGCA 481
2149 GGGGTCCGCTGTTGTGCGCCGCGGAGCGCGGTGGGCAATTTAGGGCGCGGTGTGCA 2208
482 CCCGTGAGGTGGCTAAGCGGTGCACTTATCCCGGTAGAGAGCTTAGAGACAACCATGA 541
2209 CCCGTGAGGTGGCTAAGCGGTGCACTTATCCCTGTGTGAGAGAACCTAGAGACAACCATGA 2268
542 GGTCTCCCGGTGTTCTCAGACAACCTCTCTCCACAGCAGTCCCGCAGAGTACCAGTGG 601
2269 GGTCTCCCGGTGTTCTCAGACAACCTCTCTCTCAGAGTGTGCGGAGAGCTTCCAGGTGG 2328
602 CCACCTCTGATGTCTCCACCGCGAGCGGTAAAGACACCAAGTCTCCCGCGCATACGAG 661
2329 CTACCTCTCATGTCTCCACAGGCGAGCGGCAAAAGCAAGTCTCCGCGTGCATATGAG 2388
662 CTAGGGCTCAAGAGTGTGCTGCTCAACCCCTCGTGTGCTGCAACATGGGCTTTGGTG 721
2389 CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGGCTTTGGTG 2448
722 CTTACATGTCCAAGGCCCATGGGATTTGATCTTAACATCAGGACTGGGGTGGAGCAATTA 781
2449 CTTACATGTCCAAGGCTCATGGGATCGATCTTAACATCAGAGCGGGGTGAGAACAAATTA 2508
782 CTACTGGAGCCCGATCACGTATTCCACCTACGGCAAGTCTCTTGGCGAGCGGGGTGTT 841
2509 CCACTGGAGCCCGATCACGTACTCCACCTACGGCAAGTCTCTTGGCGAGCGGGGTGCT 2568
842 CAGGGGTGCTTATGACATTAATTTGTGACAGTGCCTCAACCGGATGCAACATCCA 901
2569 CGGGGGCGCTTATGACATTAATTTGTGACAGTGCCTCAACCGGATGCAACATCCA 2628
902 TCTTGGGCAATGGCACTGTCTTGACCAAGCAGAGACCGGGGGGAGAGTGTGCTGTC 961
2629 TCTTGGGCAATGGGCACTGTCTTGACCAAGCAGAGACTGGGGGGGAGAGTGTGCTGTC 2688
962 TCGCCACCGCTACCCCTCCGGGCTCGGTCACTGTGCCCCCATCTCAACATCAGAGAGTTG 1021
2689 TCGCCACCGCCACCCCTCCGGGCTCGGTCACTGTGCCCCCATCTCAACATCAGAGAGTTG 2748
1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATTA 1081
2749 CTCTGTCCACCAACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATTA 2808
1082 AGGGGGGAGACATCTCATCTTCTGCGCTCAATGCGGCTGCTTACTACCGCGCTTGTGCTGCA 1141
2809 AGGGGGGAGACATCTCATCTTCTGCGCTCAATGCGGCTGCTTACTACCGCGCTTGTGCTGCA 2868
1142 AACTGTGCTGGGCTGGGCTCAATGCGGCTGCTTACTACCGCGCTTGTGCTGCTGCA 1201
2869 AGCTGTGCTGCAATGGGCACTCAATGCGGCTGCTTACTACCGCGCTTGTGCTGCTGCA 2928
1202 TCCGACCAAGTGTGAGTGTGCTGCTGCTGCACTGACGCGCTCTACGCGCTTTTACCG 1261
```

Db 2929 TCCGACAGCGGCGATGTTGTCGTGCGTGGCAACCGATGCCCTCATGACCGGTATACCG 2988
Qy 1262 GCAGATTCGATTCGGTATAGACTGCAACACGCTGTGTCACCCAGACAGTGCATTCAGCC 1321
Db 2989 CGGACTTCGACTCGGTGATAGACTGCAATACGTTGTGTCACCCAGACAGTGCATTCAGCC 3048
Qy 1322 TTGACCCCTACCTTCACCATTTGAGACAATACGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381
Db 3049 TTGACCCCTACCTTCACCATTTGAGACAATACGCTTCCCGAGGATGCTGTCTCCCGCACTC 3108
Qy 1382 AACGTCGGGTAGACTGCGCAGAGGGAAGCCAGGACTCTACAGATTTGTGGCAGCGGGG 1441
Db 3109 AACGTCGGGCGAGACTGCGCAGAGGGAAGCCAGGACTCTACAGATTTGTGGCAACCGGGG 3168
Qy 1442 AGGTCCTCTTCGACATGTTGACTCGTCTGCTCTCGGAGTCTATGACGCGGTTGTG 1501
Db 3169 AGGCCCCCTCCGGCATGTTGACTCGTCTGCTCTCTGTGAGTCTATGACGAGGCTGTG 3228
Qy 1502 CTTGGTATGACTTACGCCCGCGAGACCAAGTTAGGCTACGAGCATATCAATCAACACC 1561
Db 3229 CTTGGTATGACTCAGCCCGCGAGACTACAGTTAGGCTACGAGCTACATGAACACC 3288
Qy 1562 CGGCACTTCGCTGTGCGAAGCAATCTTGAAATTTTGGGAGGCGCTTTTACGGGTCTCA 1621
Db 3289 CGGGGCTTCCCGTGTGCGAGGACCATCTTGAAATTTTGGGAGGCGCTTTTACAGGCTCA 3348
Qy 1622 CCACATAGACGCCACTTCTCTATCCAGACAAGCAGAGTGGGGAAACCTTCCCTATC 1681
Db 3349 CTATATAGATGCCACTTCTATCCAGACAAGCAGAGTGGGGAAACCTTCTCTTACC 3408
Qy 1682 TGSTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCGTCTGTGGGACC 1741
Db 3409 TGSTAGCGTACCAAGCCACCGTGTGCGCTAGGCTCAAGCCCTCCCGCATCGTGGGACC 3468
Qy 1742 AGATGTGGAAGTGTGATCCGCTCTAAGCCACCTCCATGAGGCGCAACACTCTGCTAT 1801
Db 3469 AGATGTGGAAGTGTGATTCGCTCTCAAGCCACCTCCATGAGGCGCAACACTCTGCTAT 3528
Qy 1802 ATAGACTGGGCGTGTCCAGAAAGTCAAGTCAAGTCAAGCGCACCCAGTCAACAGTATATCA 1861
Db 3529 ACAGACTGGGCGTGTTCAGAAATGAAATCACTTGAAGCAGCACCAGTCAACAAATATCA 3588
Qy 1862 TGACATGTATGTGCGGTGACTTGAGTGTGCTACAGAGTACCTGGGTGCTGTGGCGGG 1921
Db 3589 TGACATGTATGTGCGCGACCTTGAGTGTGCTACAGAGCACTTGGGTGCTGTGGCGGG 3648
Qy 1922 TTCTGGTGTCTTGGCGCGGTATGCTATCCAGCTGCGTGGTGCATAGTAGTAGGA 1981
Db 3649 TCCTGGCTGTCTTGGCGCGGTATGCTGTCAACAGGCTGCGTGGTCAATAGTGGCGAGG 3708
Qy 1982 TTGCTTGTCCGAAAGCGGCAATCATACCGACAGGGAAGTCTCTACCGGAGTTCG 2041
Db 3709 TCCTTGTTCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTCG 3768
Qy 2042 ATGAAATGGAAGTGTCT 2059
Db 3769 ATGAGATGGAAGTGTCT 3786

RESULT 11
109331

LOCUS 109331 7310 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 15 from Patent WO 8904669.
ACCESSION 109331
VERSION 109331.1 GI:587966
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7310)
AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 15 01-JUN-1989;
FEATURES Location/Qualifiers

source 1. .7310
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 86.7%; Score 1786; DB 6; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
Qy 2 TGGCGCTATACGGCTATGCCAGCAGACAAAGGGGCCCTTTGGGATGCATATACCA 61
Db 1729 TGGCGCCATACGGCGTACGCCAGACAAAGGGGCCCTCTAGGGTGCATAATACCA 1788
Qy 62 GCTTGACCGCGCGGAGCAAAACCAAGGTGAGGTTCAGATCGCTCAACTGCTG 121
Db 1789 GCTTAACCTGCCCGGCAAAACCAAGTGGAGGTGAGTCCAGATTGTCAACTGCTG 1848
Qy 122 CCAGACTTTCTTGGCAACTGCAATTAACGGGGTGTGTGGACTGTCTACCATGGAGCG 181
Db 1849 CCCAAACCTTCTGGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCAAGGGCG 1908
Qy 182 GAACAGGACCAATTCGCTCACCCTAAGGTCTCTGTTATCCAGATGTACCAATGTGGAC 241
Db 1909 GAACGAGGACCAATTCGCTCACCCTAAGGTCTCTGTTATCCAGATGTATACCAATGTAGACC 1968
Qy 242 AAGACCTCGTAGCGTGGCGCGCTCCCAAGGTGCCGCTATTAAACCATGCACCTTGGC 301
Db 1969 AAGACCTGTGGCTGGCGCGCTCCCAAGGTAGCGCTATTGACACCTGCACCTTGGC 2028
Qy 302 GCTCTCGGACCTTTACCTGTGTCAGAGGACCGCGATGCTATTCCTGTGCGCGGACGG 361
Db 2029 GCTCTCGGACCTTTACCTGTGTCAGAGGACCGCGATGCTATTCCTGTGCGCGGCG 2088
Qy 362 GTGATGGAGGGGAGCGCTGCTTTCGCCCGCGGCTATCTCTTACTTGAAGGCTCTCGG 421
Db 2089 GTGATGAGGAGGGGAGCGCTGCTTTCGCCCGCGGCTATCTCTTACTTGAAGGCTCTCGG 2148
Qy 422 GAGGCCCTCTGCTGTGCGCGGAGGACATGCGTGGGCAATTTACAGAGCGCGGTATGCA 481
Db 2149 GGGGTCCGCTGTGTCGCCCGGGGACCGCGTGGGCAATTTAGGGCGCGGTGTGCA 2208
Qy 482 CCCGTGGAGTGTCTAAGCGGTGGAATTCATCCCGTAGAGACTTTAGAGCAACCATCA 541
Db 2209 CCCGTGGAGTGTCTAAGCGGTGGAATTTATCCCTGTGGAGAACCTTAGAGCAACCATCA 2268
Qy 542 GGTCCCGGTGTCTCAGACAACTCCTCCCAAGCAGCTGCCCCAGAGCTACCAAGTGG 601
Db 2269 GGTCCCGGTGTCTCAGGATAACTCCTCTCCCAAGTGTGCCCCAGAGCTTCCAGGTGG 2328
Qy 602 CCCACTGTCATGCTCCCAAGCGGTAAAGGCAACCAAGTCCCGGCGCATAGCGAG 661
Db 2329 CTCACCTCATGCTCCCAAGCGGTAAAGGCAACCAAGTCCCGGCGCATATGCAAG 2388
Qy 662 CTCAGGGCTACAAGGTGTGCTCAACCCCTCCGTTGCTGCTCAACAAATGGGCTTTGGTG 721
Db 2389 CTCAGGGCTATAAGGTGTGCTCAACCCCTCTGTTGCTCAACACTGGGCTTTGGTG 2448
Qy 722 CTTACATGTCGAAGCCCATGGGATTTGATCTTAAATCATAGGACTGGGTGAGGACAAATTA 781
Db 2449 CTTACATGTCGAAGGCTCATGGGATTCGATCTTAAATCATAGGACCGGGGTGAGGACAAATTA 2508
Qy 782 CTACTGGGAGCCCGATCAGCTATTCACCTACGGCAAGTTCCTTCCGACGGCGGTGT 841
Db 2509 CCACCTGGGAGCCCGATCAGCTACTCCACTAGCGCAAGTTCCTTCCGACGGCGGTGT 2568
Qy 842 CAGGGGTGCTTATGACATAATAATTTGTGACGAGTGGCACTCCACGGATGCAACATCA 901
Db 2569 CAGGGGTGCTTATGACATAATAATTTGTGACGAGTGGCACTCCACGGATGCAACATCA 2628
Qy 902 TCTTGGGCAATTTGGCACTGCTCTTGACCAAGCAGACCGCGGGGCGAGCTGACTGTGC 961
Db 2629 TCTTGGGCAATTTGGCACTGCTCTTGACCAAGCAGACCTGCGGGGCGAGCTGCTGTGC 2688

```
Qy 962 TCGCCACCGGTACCCCTCCGGGCTCGGTCACTGTGCCCCCATCTCTAAATCATGAGAGGTTG 1021
Db 2689 TCGCCACCGCACCCCTCCGGGCTCGGTCACTGTGCCCCCATCTCTAAATCATGAGAGGTTG 2748
Qy 1022 CTCTGTCCATACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAACAATTA 1081
Db 2749 CTCTGTCCATACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAAGTAATCA 2808
Qy 1082 AGGGGGGAGACATCTCTATCTTTCGCACTCAAAAGAAAGTGCAGAGCTCGCCGCAA 1141
Db. 2809 AGGGGGGAGACATCTCTATCTTTCGCACTCAAAAGAAAGTGCAGAGCTCGCCGCAA 2868
Qy 1142 AACTGTCGCTTGGGCGTCAATCGCTGCTTACTACCGGCGCTTGATGTGCGCTCA 1201
Db 2869 AGTGTGCGATTTGGGCGTCAATCGCTGCTTACTACCGGCGCTTGATGTGCGCTCA 2928
Qy 1202 TCCGACACGATGATGCTGTGCTGCTGCAACTGACGCGCTCATGACGCGCTTACCG 1261
Db 2929 TCCGACACGCGGATGCTGTGCTGCTGCAACCGATGCCCTCATGACGCGCTTACCG 2988
Qy 1262 GCGACTTCGATTCGATGACTGCAACACGCTGTGTCAACGACAGACTGCACTTCAGCC 1321
Db 2989 GCGACTTCGATTCGATGACTGCAATACGCTGTGTCAACGACAGACTGCAATTCAGCC 3048
Qy 1322 TTGACCTTACCTTACCATTTAGACATCAAGCTTCCCGAGGATGCTGTCTCCGTTACTC 1381
Db 3049 TTGACCTTACCTTACCATTTAGACATCAAGCTTCCCGAGGATGCTGTCTCCCGACTC 3108
Qy 1382 AACGTCGGGTAGGACTGCGAGAGGGAAGCAGGCACTACAGATTTGTGCGACCGGGG 1441
Db 3109 AACGTCGGGAGGACTGCGAGAGGGAAGCAGGCACTACAGATTTGTGCGACCGGGG 3168
Qy 1442 AGCTTCCTTCTGCGCATGTTTGAATCTGCTGCTCTGCGAGTGTATGACGCGGTTGTG 1501
Db 3169 AGCGCCCTCCGCGCATGTTGACTCGTCTGCTCTGCTGAGTGTATGACGCGGCTGTG 3228
Qy 1502 CTTGATGATGACTTACCGCCGCGAGACCAAGTATGAGTGTACGAGCATATCAAGAACCC 1561
Db 3229 CTTGATGATGACTTACCGCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACCC 3288
Qy 1562 CGGACTTCCGCTGTCGCAAGACCATCTGAAATTTTGGGAGGCGTCTTACGCGTCTCA 1621
Db 3289 CGGAGCTTCCGCTGTCGCAAGACCATCTGAAATTTTGGGAGGCGTCTTACGCGCTCA 3348
Qy 1622 CCCACATAGACGCCACTTCTCTATCCAGACAAAGCAGAGTGGGAAAACTTCCCTTATC 1681
Db 3349 CTATATAGATGCCACTTCTCTATCCAGACAAAGCAGAGTGGGAAAACTTCTCTTACC 3408
Qy 1682 TGGTAGGCTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTTCCCGCTGCTGGGACC 1741
Db 3409 TGGTAGGCTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTTCCCGCTGCTGGGACC 3468
Qy 1742 AGATGTGGAAGTCTTGCATCCGCTCAAGCCACCTTCCATGGCCCAACACCTCTGCTAT 1801
Db 3469 AGATGTGGAAGTCTTGCATCCGCTCAAGCCACCTTCCATGGCCCAACACCTCTGCTAT 3528
Qy 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCACTCCCTGACGCAACCCAGTACCAAGTATCA 1861
Db 3529 ACAGACTGGGCGCTGTTCAGAAATGAAGTCACTCCCTGACGCAACCCAGTACCAAGTATCA 3588
Qy 1862 TGACATGATGTGCGGTGACCTGGAGTGTGCTACAGAGTACTGGGTGCTCTGTTGGCGGG 1921
Db 3589 TGACATGATGTGCGGTGACCTGGAGTGTGCTACAGAGTACTGGGTGCTCTGTTGGCGGG 3648
Qy 1922 TTCTGGCTGCTTGGCGCGGTATTTGCTTATCCACAGGCTGCTGCTATAGTAGTAGGA 1981
Db 3649 TCCTGGCTGCTTGGCGCGGTATTTGCTTATCCACAGGCTGCTGCTATAGTAGTAGGA 3708
Qy 1982 TTGCTTGTGCGGAAACCGCGCAATCATACCCGACGGAAGTCTCTTACCGGAGTTCG 2041
Db 3709 TCGTCTTGTGCGGAAACCGCGCAATCATACCCGACGGAAGTCTCTTACCGAGAGTTCG 3768
Qy 2042 ATGAAATGGAAGAGTGTCT 2059
```

```
Db 3769 ATGAGATGGAAGAGTGTCT 3786
```

RESULT 12 HPCPOLYP

LOCUS
DEFINITION
Hepatitis C virus polyprotein gene, partial cds.
ACCESSION
M32084
VERSION
M32084.1
KEYWORDS
polyprotein.
SOURCE
Hepatitis C virus
ORGANISM
Hepatitis C virus

7310 bp ss-RNA linear VRL 02-AUG-1993

REFERENCE
AUTHORS
Choo, Q.-L., Richman, K. and Han, J.
TITLE
The nucleotide sequence of the Hepatitis C viral genome
JOURNAL
Unpublished (1990)
COMMENT
Original source text: Hepatitis C virus, cDNA to viral RNA, clones K9-1 through 15e, isolated from chimpanzee (individual 910) blood plasma.
Draft entry and printed sequence for [1] kindly submitted by M. Houghton, 22-FEB-1990. Chiron Corporation, 4560 Horton Street, Emeryville CA 94608.

FEATURES
source

Location/Qualifiers
1. 7310
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/db_xrefs="taxon:11103"
<1..57310
/note="polyprotein"
/codon_start=3
/protein_id="AAA45677.1"
/db_xref="GI:329876"
/translation="GCPERLASCRPLTDDQGWGPIYSANGSDQDPQYPCWHPKPC
GIVPAKSCGVPYCTFTSPVVTGRTDRSGAPTSWGENDDVFLNNRPLGNWFGC
TWNSTGFTKVCAGPCVIGGAGNNTLHCPTDCFRKHPDATYSCSGSPITPRCLVD
YPRLMWHPCTINITYFKIRMYGVGHRLEAACNWRGRCDCDRSGLSPLLLT
TQWQVLPCTFTLPALSTGLIHLHONIVDVQYLVGSSIASWAKWYVLLFLLL
ADARVCSCLMMLLSQAEALNVLINAAAGTGLVSLVFCFANVKGKWP
GAYTYGHWPLLLLLLALPORAYALDTEVAASCQGVVLGLMALILSPYKRIKSWC
LWMLQFLTRVRAQLHVPPLNVRGRDAILLMCAVHTPLVDITKLALVAGPLW
LQASLLKVPYFVRVQGLLFCALARKMIGGHVQVMIKLGALTYGVNHLPLRD
WAHNGRDLAAVAVPVSQMETKLTWAGDTAAACGDIINGLPSVARGREIILGPAD
GMVSKWRLLAPITAYAOQTGILGCIITSLTGRDKNOVEGEVQIIVSTAQTPLATCI
NGVCTVYHGAGTTRTIASPKGVIMYTNVDQDLGVMPAQSGRSILTPCTCGSDLYL
KTHADVIPIRRRGRSGRLSPISYLGKSGGFLLCFAGHAGVIFRAAVCTGVA
KAVDFIPVENLETTMRSPVFTDNPSPVVPVQSFQVAHLHAPTGGSKSTRVPAAYAQG
YKVLNPSVAATLGFAYMSKAHGIDPNIRTVGRTITGTSPTSTYTGKFLADGGGS
GGAYDIIICDECHSTDATSLIGTGLDQAEATAGARLVLATATPGSVTVPHNIEE
VALSTTGEIPIFYGKAIPLVYKGRHLIFCHSKKCDLAAKVALGINAVAYRGUD
VSVITPGEIVVAVDALMTGYTGDFSDVIDCNTQTVTDVFSLDPTFTTETLPOD
AVSRTORRGTRGKKGITRFAVAGERPSPGMSVLCEDYDAGCAWELTAPETTVR
LRAYNMTGPLPVQCQDHLFEWGTGLTHIDAHFLSQTKQSGENLPYLVAQATVCAR
AQAPPSQWQMKLLRIKPLTLPGLTPTLLYRLGAVQNEITLTHPTVYIMTCDMSADLE
WHTLWLVGVLAAALAAACLSGCVIVGVVVLGSKPAIIPDREVLVYRFDMEBES
OHLPTYEQGMIAEQFKOKALGLLOTASRAQAEVIAPAVOTNWKLETFFMAKHMNFTS
GTQYLAGLSTLPCNPALASLMAFTAVTSPITTSOTILLFNILGWAAQAALAPGAATA
FVAGLAGAAGSVGKGLVIDILAGTAGAGVAGALVAFKIMSGSEVSTEDLNLPLAI
LSPGALVGVCAAILRRHVPGEGAVQWNNRLIAFASRGNHVSPTHYPESDAARV
TLLISLUTVQLLRHQMISSECTTCCSGWLRIDWMTCEVLSDFKTLWKLAKMGP
LPGIPVSCORGKGVRVGDIHMTTRCGCAEITGHVXNGTMRIVGRTCRNWMSTQF
PINAVTGCTPLPAPNYTFAWRVSAEEVEIROVGDPHVTGMTDNLKPCQVPS
PEPFTLGDVRLHRPAPCKPLLRBSVSRVGLHYPVGSQLEPCPEBPDAVIELSMLT
DPSHITAEAGRLRLARGSPSVASSASQLSKATCTANHDSRDEADIELNLMWR
QEMGNITVSEKENVILDSFDPLVAEEDEREISVPAETLRLKRSRFAQALPWAPRD
YNPLPVTWKPKDYPPEPVHGCPLPPPKPPVPPVPPKRTVLTTESTLSTALAEATR
SFSSGTSSTIGDNTTSSPEPAPCPDSDAESYSMPLEGEPPGDPDLDSLSWSV
SEANAEADVCCSMYSYTCALVTPCAAEQKLPINALSNLHNLHNLVSTYTSRAC
QRQKQYFRLQVLDHYQDVLKEVAAASKVKANLLSVEEACSLIPPHSKSKFGVG
AKDVRCHARKAVTHINSWKDLEDNVTPIDITIMAKNEVFCVQPEKGRKPARLIVF
PDLGVRVCEKMAIYDVVTKLPLAVMGSSYGFQYSPQGRVEFLVQAWKSKKTPMGFSYD

TRCPDSTVTSIDIRTEBAIYQCCDLDPQARVAIKSLTERLYVGGPLNTRSGENCGYRR
CRASGVLTSCGNLTCTCIKARAACRAAGLQDCTMLVCGDDLVIVICESAGVQBDASL
RAFTEAMTRYSAPEPDPQPEYDLELITSCSSNVSAHDAKGKRVYLTRDPTPLAR
AAWETARHTVPUNSHLNIIMFAPILMARMILMTHFFSVLIARDQLEQALDCEIYGACY
SIEPLDLPPIIQRLL"

ORIGIN

Query Match 86.7%; Score 1786; DB 14; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy	2	TGGGCGCTATCAGCGCCTATGCCCAGCAGACAAGGGGCGCTTTGGGATGATAATCACCA	61
Db	1729	TGGGCGCCATCAGCGCCTAGCCCCAGCAGACAAGGGGCGCTTCTAGGGTGATAATCACCA	1788
Qy	62	GCTTGACCGCGCGGACAAACACAGGTGGAGGGTGAGGTTCCAGATCGTGTCAACTGCTG	121
Db	1789	GCCTAACTGGCGCGGACAAACACAGGTGGAGGGTGAGGTTCCAGATGCTGTCAACTGCTG	1848
Qy	122	CCAGAGCTTTCTTGGCAACCTGCATTAACCGGGTGCTGTGCACTGTCTACCATGGAGCGG	181
Db	1849	CCCAAACTTCTTGGCAACCTGCATTAACCGGGTGCTGTGCACTGTCTACCATGGAGCGG	1908
Qy	182	GAACAAGGACCATTTGCGTCACTTAAGGGTCTGTGTTATCCAGATGTACACCAATGTGAGCC	241
Db	1909	GAACGAGGACCATCGCGTCACTTAAGGGTCTGTGTTATCCAGATGTATACCAATGTAGACC	1968
Qy	242	AAGACCTCGTAGGTGCGCGCTCCCAAGGTGCGCGCTCAITTAACCAATGCACTTGCG	301
Db	1969	AAGACCTGTGGGTGCGCGCTCCCAAGGTGCGCGCTCAITTAACCAATGCACTTGCG	2028
Qy	302	GCTCTCGGACCTTTACTCTGTGTCACAGGACGCGCGATGTCATTCTGTGCGCGGACGGG	361
Db	2029	GCTCTCGGACCTTTACTCTGTGTCACAGGACGCGCGATGTCATTCTGTGCGCGGCGGG	2088
Qy	362	GTATGACGAGGGAGCGCTGCTTTGCGCCCGCGCTATCTCTTAATGAAAGGCTCTCTCGG	421
Db	2089	GTATAGCAGGGGAGCGCTGCTGTCGCGCGCGCGCTATCTCTTAATGAAAGGCTCTCTCGG	2148
Qy	422	GAGCGCTCTGCTGTGCGCGGAGGACATGCGGTAGGCAATATTCAGAGCGCGGATGCA	481
Db	2149	GGGGTCCGCTGTGTGCGCGCGGAGGACATGCGGTAGGCAATATTTAGGCGCGGATGCA	2208
Qy	482	CCCGTGGAGTGGCTAAGCGCGGTGGACTTCAATCCCGGTAGAGAGCTTAGAGACAACATGA	541
Db	2209	CCCGTGGAGTGGCTAAGCGCGGTGGACTTATCCCTGTGGAGAACCTTAGAGACAACATGA	2268
Qy	542	GGTCCCGGTGTTCTCAGACAATCTCTCCCGCAGAGTGCGCCAGAGCTACCAAGTGG	601
Db	2269	GGTCCCGGTGTTCTCAGGATAACTCTCTCCACAGTAGTGGCCCGCAGAGCTTCCAGGTGG	2328
Qy	602	CCCACTGCAATGCTCCCGCGGAGCGGTAAAGGACCAAGGTCCCGCGCGCATACGCAG	661
Db	2329	CTCACTCCATGCTCCCAAGGAGCGGCAAGAGCAACAGGTCCCGCGGTGCATATGCAG	2388
Qy	662	CTCAGGGCTACAAGGTGCTGTGCTCAACCCCTCGGTGCTGCAACAATGGGCTTTGGTG	721
Db	2389	CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGCTTTGGTG	2448
Qy	722	CTTACATGTCGAAGGCCATGGGATGATCTTAACATCAGGACTGGGGTGAGGACAATTA	781
Db	2449	CTTACATGTCGAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA	2508
Qy	782	CTACTGGCAGCCCGATCAGCTATTTCCACTACGGCAAGTTCTCTTGGCGAGCGGGGTGT	841
Db	2509	CCACTGGCAGCCCGATCAGCTATTTCCACTACGGCAAGTTCTCTTGGCGAGCGGGGTGT	2568
Qy	842	CAGGGGTGCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAACATCCA	901
Db	2569	CGGGGGCGCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCCATCCA	2628
Qy	902	TCTTGGGCATTGGCACTGCTTGTGACCAAGCAGAGACCGGGGGGCGAGCTGACTGTGC	961
Db			

Db	2629	TCCTTGGGCATCGSCACTGTCTCTTGACCAAGCAGAGACTCGGGGGCGAGACTCGTTGTGC	2688
Qy	962	TGCCACACGCTACCCCTCGGGGCTCGTCACTGTGCCCCCATCTTAACATCGAGGAGTTG	1021
Db	2689	TGCCACACGSCACCCCTCGGGGCTCGTCACTGTGCCCCCATCTTAACATCGAGGAGTTG	2748
Qy	1022	CTCTGTCCCACTACCGGAGAGATCCCTCTTTATGGAAGGCTATTCCCTTTGAAGCAATTA	1081
Db	2749	CTCTGTCCCACTACCGGAGAGATCCCTCTTTATGGAAGGCTATTCCCTCTCGAAGTAATCA	2808
Qy	1082	AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTGCGACGAGCTCGCCGCAA	1141
Db	2809	AGGGGGGAGACATCTCATCTTCTGTCACTCAAAGAAGAGTGCGACGAACTCGCCGCAA	2868
Qy	1142	AACTGTCGCGGTGGCGCTCAATGCGGTGCTTACTACCGCGGCTTGATGTCTCGTCA	1201
Db	2869	AGCTGTGTCATTTGGGCGCATCAATGCGGTGCTTACTACCGCGGCTTTGACGTGTCTCGTCA	2928
Qy	1202	TCCGACACAGTGTGACGTTGTGCTGTGGCAACTGACGCCCTCATGACCGGCTTTACCG	1261
Db	2929	TCCGACACAGCGGCGATGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGCTATACCG	2988
Qy	1262	GCGACTTCGACTCGGTGATAGACTGCAATACGTTGTGTCAACCAGACAGTCGACTTCAGCC	1321
Db	2989	GCGACTTCGACTCGGTGATAGACTGCAATACGTTGTGTCAACCAGACAGTCGACTTCAGCC	3048
Qy	1322	TTGACCCCTACCTTCAACATTGAGACATCACGCTTCCCAGGATGCTGTCTCCGCTACTC	1381
Db	3049	TTGACCCCTACCTTCAACATTGAGACATCACGCTTCCCAGGATGCTGTCTCCGCTACTC	3108
Qy	1382	AACGTCGGGGTAGGACTGCGCAGAGGAGACAGGATCTTACAGATTTGTGGCAACCGGGG	1441
Db	3109	AACGTCGGGGTAGGACTGCGCAGAGGAGACAGGATCTTACAGATTTGTGGCAACCGGGG	3168
Qy	1442	AGCGTCTTCTTGGCATGTTGTGACTCGTCTGTCTCTGCGAGTGCTATGACGCGGTTGTG	1501
Db	3169	AGCGTCTTCTTGGCATGTTGTGACTCGTCTGTCTCTGCGAGTGCTATGACGCGGTTGTG	3228
Qy	1502	CTTGTGTATGAGCTTACGCCCGCGAGACACAGTTAGGCTACGACATACATGACACACC	1561
Db	3229	CTTGTGTATGAGCTTACGCCCGCGAGACACAGTTAGGCTACGACATACATGACACACC	3288
Qy	1562	CGGGACTTCTCCCGTGTGCCAAGACCATCTTTGAATTTTGGAGGCGCTCTTTACGGGTCTCA	1621
Db	3289	CGGGCTTCTCCCGTGTGCCAAGACCATCTTTGAATTTTGGAGGCGCTCTTTACGGGTCTCA	3348
Qy	1622	CCCATAGACGCCCATCTTCTATCCAGACAAAGAGAGTGGGAAACCTTCCCTATC	1681
Db	3349	CTCATATAGATCCCACTTCTTATCCAGACAAAGAGAGTGGGAAACCTTCCCTATC	3408
Qy	1682	TGCTAGCTTACCAAGCCCGTGTGCTTAGAGCTCAAGCCCTCCCGTCTCGTGGGACC	1741
Db	3409	TGCTAGCTTACCAAGCCCGTGTGCTTAGAGCTCAAGCCCTCCCGTCTCGTGGGACC	3468
Qy	1742	AGATGTGAAGTGTGTGATCGCTCTCAAGCCCAACCTCCATGGGCAACACCTCTGCTAT	1801
Db	3469	AGATGTGAAGTGTGTGATCGCTCTCAAGCCCAACCTCCATGGGCAACACCTCTGCTAT	3528
Qy	1802	ATAGACTGGGCGCTGTCCAGATGAAAGTCAACCTGACGCAACCTGACCAAGTATATCA	1861
Db	3529	ACAGACTGGGCGCTGTCCAGATGAAAGTCAACCTGACGCAACCTGACCAAGTATATCA	3588
Qy	1862	TGACATGTATGTGCGGCTGACCTGGAGGTGCTCACAGTACCTGGGTGCTCGTTGGCGGG	1921
Db	3589	TGACATGTATGTGCGGCGACCTGGAGGTGCTCACAGTACCTGGGTGCTCGTTGGCGGG	3648
Qy	1922	TTCTGGCTGCTTTGGCCCGCTATTGCTCTATCCAGGCTGCGTGTCTATGATGAGTAGGA	1981
Db	3649	TTCTGGCTGCTTTGGCCCGCTATTGCTCTATCCAGGCTGCGTGTCTATGATGAGTAGGA	3708
Qy	1982	TTGTCTTGTCCGGAAGCGGCAATCATACCGACAGAGGAGTCTCTTACCGGGAGTTGCG	2041
Db	3709	TGCTCTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTGCG	3768

Qy 2042 ATGAAATGGAAGAGTGCT 2059
| | | | |
Db 3769 ATGAGATGGAAGAGTGCT 3786

RESULT 13
108294
LOCUS 108294 9185 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent EP 0388232.
ACCESSION 108294
VERSION 108294.1 GI:588994
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9185)
AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.
TITLE NABV diagnostics and vaccines
JOURNAL Patent: EP 0388232-A1 1 19-SEP-1990;
FEATURES
Location/Qualifiers
source 1. 9185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 86.7%; Score 1786; DB 6; Length 9185;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGGGCTATACGGCTATGCCAGAGACCAAGGGGCTTTGGGATGATATACCA 61
| | | | |
Db 3396 TGGGGCCCATCACGGCTATGCCAGAGACCAAGGGGCTTCTAGGGTGCATATACCA 3455

Qy 62 GCTTGACCGCGCGGACAAACAGGTGGAGGTGAGTTGAGTCTGCTCAACTGCTG 121
| | | | |
Db 3456 GCCTAACTGGCGCGGACAAACAGGTGGAGGTGAGTTGAGTCTGCTCAACTGCTG 3515

Qy 122 CCCAGACTTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCGG 181
| | | | |
Db 3516 CCCAACTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCGG 3575

Qy 182 GAACAAGACCATTTGCGTACCTAAGGTCTCTGTTATCCAGATGATACCAATGTGAAC 241
| | | | |
Db 3576 GAACAAGACCATTTGCGTACCTAAGGTCTCTGTTATCCAGATGATATACCAATGTGAAC 3635

Qy 242 AAGACCTCGTAGGCTGCGCGCTCCCAAGGTGCGCTCATTAACACCATGCACTTGG 301
| | | | |
Db 3636 AAGACCTCGTAGGCTGCGCGCTCCCAAGGTGCGCTCATTAACACCATGCACTTGG 3695

Qy 302 GCTCTCGGACCTTTACTGTGTACAGGACGCGGATGTCATTCCTGTGCGCGGACGGG 361
| | | | |
Db 3696 GCTCTCGGACCTTTACTGTGTACAGGACGCGGATGTCATTCCTGTGCGCGGACGGG 3755

Qy 362 GTGATGACGGGCGAGCTGCTTTGCGCGCGCTATCTCTTACTTGAAGGCTCTCGG 421
| | | | |
Db 3756 GTGATGACGGGCGAGCTGCTTTGCGCGCGCTATCTCTTACTTGAAGGCTCTCTCGG 3815

Qy 422 GAGCCCTCTGCTGCGCGCGGACGACNTCCGTAGGCATATTCAGAGCCGCGTATGCA 481
| | | | |
Db 3816 GAGCCCTCTGCTGCGCGCGGACGACNTCCGTAGGCATATTCAGAGCCGCGTATGCA 3875

Qy 482 CCCGTGAGTGGCTTAAGCGGTGACTTCTATCCCGGTAGAGAGCTTAGAGACCAACATGA 541
| | | | |
Db 3876 CCCGTGAGTGGCTTAAGCGGTGACTTCTATCCCGGTAGAGAGCTTAGAGACCAACATGA 3935

Qy 542 GGTCCCGGTGTTCTCAGACAACTCTCCCAACAGCAGTGCCTCAGAGCTTACCAAGTGG 601
| | | | |
Db 3936 GGTCCCGGTGTTCTCAGACAACTCTCTCCACAGTAGTGCCTCAGAGCTTCCAGGTGG 3995

Qy 602 CCACCTGCTATGCTCCCAACGGGAGCGGTAGAGCAACAGGTCCCGGCGCATACGAG 661
| | | | |
Db 3996 CTCACTCCATGCTCCCAACGGGAGCGGTAGAGCAACAGGTCCCGGCGCATACGAG 4055

Qy 662 CTCAGGGCTACAAAGTGTGCTCAACCCCTCGTTGCTGCAACAATGGCTTTGGTG 721
| | | | |
Db 4056 CTCAGGGCTATAAGTGTGCTAGTACTCAACCCCTCTGTTGCTGCAACAATGGCTTTGGTG 4115

Qy 722 CTTACATGTCCAAGGCCCATGGGATGATCTTAACATCAGAGCTGGGGTGGAGCAATTA 781
| | | | |
Db 4116 CTTACATGTCCAAGGGCTCATGGGATCGATCTTAACATCAGAGCCGGGTGAGAACAAATTA 4175

Qy 782 CTACTGCGACCCGATCAACGTATTCCACCTACGCAAGTTCTTTCGCGACGGCGGTGTT 841
| | | | |
Db 4176 CCACTGGCAGCCCATCACGTACTCCACCTACGGCAAGTTCTTTCGCGACGGCGGTGCT 4235

Qy 842 CAGGGGTGCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCACATCCA 901
| | | | |
Db 4236 CGGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCCACATCCA 4295

Qy 902 TCTTGGCATTTGGCATGTCTCTTGACCAAGCAGAGACCGGGGGCGAGACTGACTGTGC 961
| | | | |
Db 4296 TCTTGGCATTTGGCATGTCTCTTGACCAAGCAGAGACTGGGGGGCGAGACTGACTGTGC 4355

Qy 962 TCGCCACCGCTACCCCTCCGGGCTCCGTCACGTGCCCCCATCTTAACATCGAGGAGTTG 1021
| | | | |
Db 4356 TCGCCACCGCACCCCTCCGGGCTCCGTCACGTGCCCCCATCTTAACATCGAGGAGTTG 4415

Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGCGCAAGGCTATTCCCTTTGAAGCAATTA 1081
| | | | |
Db 4416 CTCTGTCCACTACCGGAGAGATCCCTTTTATGCGCAAGGCTATTCCCTTTGAAGCAATTA 4475

Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAAGTGGGAGAGCTGCGCGCAA 1141
| | | | |
Db 4476 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAAGTGGGAGAGCTGCGCGCAA 4535

Qy 1142 AACTGGTTCGGTGGGGCTCAATGCGTGGCTTACTACCGCGCTTGATGTGCGGTCA 1201
| | | | |
Db 4536 AACTGGTTCGGTGGGGCTCAATGCGTGGCTTACTACCGCGCTTGATGTGCGGTCA 4595

Qy 1202 TCCGACACAGTGTGAGTGTGCTGCTGCGCAACTGACGCGCTCATGACCGGCTTTACCG 1261
| | | | |
Db 4596 TCCGACACAGTGTGAGTGTGCTGCTGCGCAACTGACGCGCTCATGACCGGCTTTACCG 4655

Qy 1262 GCACCTTCGATTCGGTGTATGACTGCACAGCTGTGTGTCACCCAGAGAGTGCATTCAGCC 1321
| | | | |
Db 4656 GCACCTTCGATTCGGTGTATGACTGCACAGCTGTGTGTCACCCAGAGAGTGCATTCAGCC 4715

Qy 1322 TTGACCCCTACTCTCACCATTTGACACATCACGCTTCCCGAGGATGCTGTCCCGTACTC 1381
| | | | |
Db 4716 TTGACCCCTACTCTCACCATTTGACACATCACGCTTCCCGAGGATGCTGTCCCGTACTC 4775

Qy 1382 AACGTGGGGTAGGACTGGCAGAGGAAAGCCAGGCATCTACAGATTTGTGGCACCGGGGG 1441
| | | | |
Db 4776 AACGTGGGGTAGGACTGGCAGAGGAAAGCCAGGCATCTACAGATTTGTGGCACCGGGGG 4835

Qy 1442 AGCGTCTTCTGGGATGTTTGAATGCTGTCTCTGCGAGTGTGATGACGCGGGTGTG 1501
| | | | |
Db 4836 AGCGTCTTCTGGGATGTTTGAATGCTGTCTCTGCGAGTGTGATGACGCGGGTGTG 4895

Qy 1502 CTTTGTATGAGTCTACCGCGCGAGACACAGTTAGGCTACGAGCATACATGAACACC 1561
| | | | |
Db 4896 CTTTGTATGAGTCTACCGCGCGAGACACAGTTAGGCTACGAGCATACATGAACACC 4955

Qy 1562 CGGGACTTCCCGTGTGCCAAGACCATCTTGAAATTTTGGGAGGGCGTCTTTACGGGTCTCA 1621
| | | | |
Db 4956 CGGGACTTCCCGTGTGCCAAGACCATCTTGAAATTTTGGGAGGGCGTCTTTACGGGTCTCA 5015

Qy 1622,CCCAATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAAAAAAGCTTCCCTATC 1681
| | | | |
Db 5016 CTATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAGAAAGCTTCTTACCC 5075

Qy 1682 TGGTAGCGTACCAAGCCCGTGTGCGTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
| | | | |
Db 5076 TGGTAGCGTACCAAGCCCGTGTGCGTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 5135


```
QY 1742 AGATGTGAAGTCTGTGATCGCTCTCAAGCCACCCCTCCATGGGCCAACACCTCTGCTAT 1801
Db 5136 AGATGTGAAGTCTGTGATTCGCTCAAGCCACCCCTCCATGGGCCAACACCCCTGCTAT 5195
QY 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTTGACGACCCAGTCAACAAGTATATCA 1861
Db 5196 ACAGACTGGGCGCTGTTCAGAAATGAAGTCAACCTTGACGACCCAGTCAACAATATATCA 5255
QY 1862 TGACATGTATCTGGCTGACCTGAGGTGCGTCAGAGTACCTGGTGTCTGTTGGCGCG 1921
Db 5256 TGACATGTATCTGGCGACCTGAGGTGCGTCAGAGTACCTGGTGTCTGTTGGCGCG 5315
QY 1922 TTCTGGTGTCTTTGGCGCGTATTTGCCCTATCCACAGCTGCGTGTCTATAGTAGGTAGGA 1981
Db 5316 TCCTGGTGTCTTTGGCGCGTATTTGCCCTATCCACAGCTGCGTGTCTATAGTAGGTAGGA 5375
QY 1982 TTGTCTTGTCCGAAAGCCGCAATATACCCGACAGGAGTCTCTACCGGAGTTG 2041
Db 5376 TCGTCTTGTCCGGAAGCCGCAATATACCTGACAGGAGTCTCTACCGAGAGTTG 5435
QY 2042 ATGAATGGAAGTGTCT 2059
Db 5436 ATGAGATGGAAGTGTCT 5453

RESULT 14
BD091382 9185 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION HCV cultivation method in eucaryotic cells.
ACCESSION BD091382
VERSION BD091382.1 GI:22636993
KEYWORDS JP 2001314192-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 9185)
Weiner,A.J., Steimer,K.S. and Houghton,M.
HCV cultivation method in eucaryotic cells
Patent: JP 2001314192-A 3 13-NOV-2001;
CHIRON CORP
OS Homo sapiens (human)
PN JP 2001314192-A/3
PD 13-NOV-2001
PF 15-MAR-2001 JP 2001075114
PR 25-AUG-1989 US 398667
PI AMY J WEINER,KATHELYN S STEIMER,MICHAEL HOUGHTON PC
C12N15/09,C12N5/10,C12N7/00//C12N7/00,C12R1.93,C12N15/00, PC
C12N5/00
CC HCV cultivation method in eucaryotic cells
FH Key Location/Qualifiers
FT source 1..9185
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..9185
/organism='Homo sapiens'
/mol type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN

Query Match 86.7%; Score 1786; DB 6; Length 9185;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTATCAGCGCTATGCCAGCAGACAAGGGCGCTTTGGGATGCAATATCA 61
Db 3396 TGGCGCGCCATCAGCGCGTACGCCAGCAGACAAGGGCGCTCTTAGGGTGCATAATCA 3455
QY 62 GCTTGAACCGCGGGAACAAACACAGTGGAGGTTCAGATCGTGTCAACTGCTG 121
Db 3456 GCCTAAGTGGCGGGAACAAACACAGTGGAGGTTCAGATCGTGTCAACTGCTG 3515
QY 122 CCAGACTTTTCTGGCAACCTGCATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCG 181
```

```
Db 3516 CCACAACCTTCTGGGCAAGTGCATCAATGGGGTGTCTGGACTGTCTACCAACGGGCGG 3575
QY 182 GAACAAGGACCAATTCGCTCACTTAAGGTCCTGTTATCCAGATGTACACCAATGTGGACC 241
Db 3576 GAACAGGAGCAATCGCGTCAACCAAGGTCCTGTATCCAGATGTATACCAATGTAGACC 3635
QY 242 AAGACCTCGTAGGTCGGCCGCTCCCAAGGTGCGGCTCATTAACACATGCACTTGGC 301
Db 3636 AAGACCTTGTGGGCTGGCCGCTCCGCAAGGTAGCCGCTCATTTGACACACCTGCACCTGG 3695
QY 302 GCTCTCCGACCTTTACTTGTGTCACGAGGACGCGGATGTCAATTCCTGTGCGCGGACG 361
Db 3696 GCTCTCCGACCTTTACTTGTGTCACGAGGACGCGGATGTCAATTCCTGTGCGCGGCGG 3755
QY 362 GTGATGGCAGGGGAGCCTGCTTTTGGCCCGGCTATCTCTTACTTTGAAGGCTCTCTCG 421
Db 3756 GTGATAGCAGGGGAGCCTGCTGTGCGCCCGGCCCATTTTCTTACTTTGAAGGCTCTCTCG 3815
QY 422 GAGGCCCTCTGCTGTGTCGCGGACGACATGCGTAGGATATTCAGAGCCGCGGTATGCA 481
Db 3816 GGGGTCGCTGTGTGTCGCGGCGGACGCGCTGGGCATATTTAGGGCGCGGCTGTGCA 3875
QY 482 CCGGTGGAGTGTAAAGGCGGTGGACTTCATCCCGTAGAGACTTTAGAGACCAACCATCA 541
Db 3876 CCGGTGGAGTGTAAAGGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACCAACCATCA 3935
QY 542 GGTCCCGGCTGTTCTCAGACAACTCTCTCCACAGCAGTGTCCCGCAGAGTACCAAGTGG 601
Db 3936 GGTCCCGGCTGTTACGGATAAATCTCTCTCCACAGTAGTGTCCCGCAGAGTCTCCAGTGG 3995
QY 602 CCCACTGCATGCTCCACCGGCGAGGTAAAGACACCAAGTCCGCGCGCATAGCAG 661
Db 3996 CTCACCTCATGCTCCACAGGCGGCAAGACCAAGGTCCCGGCTGCATATGCA 4055
QY 662 CTCAGGGCTTACAAGTGTGCTCAACCCCTCGTGTCTCAACAATGGGCTTTGGTG 721
Db 4056 CTCAGGGCTTAAAGTGTGCTAGTACTCAACCCCTCTGTTGCTCAACACTGGGCTTTGGTG 4115
QY 722 CTTACATGTCCAAAGGCCAATGGGATGTATCTTAAATCAGAGACTGGGCTGAGGACAAATTA 781
Db 4116 CTTACATGTCCAAAGGCTCATGGGATCGATCTTAAATCAGAGACCGGGGTGAGAACAAATTA 4175
QY 782 CTACTGGCAGCCGATCAGTATTCACCTAGCGCAAGTCTCTCCGACGCGGGGTGT 841
Db 4176 CCAGTGGCAGCCGATCAGTACTCCACTCAGGCAAGTCTCTTCCGACGCGGGGTGT 4235
QY 842 CAGGGGCTGTCTATGACATAATAATTTGTGACGAGTGCACCTCCACGGATGCAACATCA 901
Db 4236 CGGGGGCGCTTATGACATATAATTTGTGACGAGTGCACCTCCACGGATGCAACATCA 4295
QY 902 TCTTGGGCAATTGGCACTGTCTTTGAACCAAGCAGACCGCGGGGCGGAGTGAATGCTGTC 961
Db 4296 TCTTGGGCAATCGGCACTGTCTTTGACCAAGCAGAGACTGCGGGGCGGAGACTGTTGTGC 4355
QY 962 TCGCCACCGCTACCCCTCGGGCTCCGTCACCTGTCGCCCATCTTAACATCGAGGAGTTG 1021
Db 4356 TCGCCACCGCTACCCCTCGGGCTCCGTCACCTGTCGCCCATCTTAACATCGAGGAGTTG 4415
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATTA 1081
Db 4416 CTCTGTCCACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATTA 4475
QY 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGAGCTCGCGGCA 1141
Db 4476 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGAGCTCGCGGCA 4535
QY 1142 AACTGCTCGGTTGGCGTCAATGCGGCTTACTACCGGGCTTCAATGTGTCGCTCA 1201
Db 4536 AGCTGCTCGATTTGGGCAATCAATGCGGCTTACTACCGGGCTTCAATGTGTCGCTCA 4595
QY 1202 TCCCGACAGTGGTGAACGTTGTGCTGTGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261
```


Db 4596 TCCGACACGCGGCGATGTTGTCGTGCGGCAACCGATGCCCTCATGACCGGCTATACCG 4655
Qy 1262 GCGACTTCGATTCGCGTGTGATAGACTGCAACACGTTGTGTCACCCAGACAGTGCAGTTTCAGCC 1321
Db 4656 GCGACTTCGATTCGCGTGTGATAGACTGCAATACGTTGTGTCACCCAGACAGTGCAGTTTCAGCC 4715
Qy 1322 TTGACCTTACCTTCACCATTTGAGACAATACAGCTTCCCGAGGATGCTGTCTCCCGTACTC 1391
Db 4716 TTGACCTTACCTTCACCATTTGAGACAATACAGCTTCCCGAGGATGCTGTCTCCCGACTC 4775
Qy 1382 AACGTCGCGGTPAGACTGCGCAGAGGAAGCAGGACATCTACAGATTTTGGSCACCGGGGG 1441
Db 4776 AACGTCGCGGTCAGACTGCGCAGGGAAGCAGGACATCTACAGATTTTGGSCACCGGGGG 4835
Qy 1442 AGCGTCTTTCGCGATGTTGACTCGTCTGTCCTCTGCGAGTGTATGACGCGGGTTGTG 1501
Db 4836 AGCGCTTCTCGGCGATGTTGCGACTCGTCTGTCCTCTGTGAGTGTATGAGCGAGCTGTG 4895
Qy 1502 CTTGGTATGACTTACGCCCGCGAGACCAAGTATAGGCTACGAGCATACATGAACACCC 1561
Db 4896 CTTGGTATGACTCACGCCCGCGAGACTACAGTTAGGCTACGAGCGGTACATGAACACCC 4955
Qy 1562 CGGACTTTCGCGTGTGCCAAGACCATCTTGAAATTTTGGAGGGGCTCTTTACGGGCTCTCA 1621
Db 4956 CGGGCTTTCGCGTGTGCCAAGACCATCTTGAAATTTTGGAGGGGCTCTTTACAGGCTCTCA 5015
Qy 1622 CCCATAGACGCGCACTTCTTATCCAGACAAGCAGAGTGGGGAACCTTCCCTATC 1681
Db 5016 CTATATAGATGCCCACTTCTTATCCAGACAAGCAGAGTGGGGAACCTTCCCTATC 5075
Qy 1682 TGGTAGGCTACCAAGCCACCGTGTGCGTACAGCTCAAGCCCTTCCCGGCTCGTGGGACC 1741
Db 5076 TGGTAGGCTACCAAGCCACCGTGTGCGTACAGCTCAAGCCCTTCCCGGCTCGTGGGACC 5135
Qy 1742 AGATGTGGAAGTCTTGCATCGCTCTCAAGCCACCTTCCATGGCGCAACCTCTGCTAT 1801
Db 5136 AGATGTGGAAGTCTTGCATCGCTCTCAAGCCACCTTCCATGGCGCAACCTCTGCTAT 5195
Qy 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAAGCTGAGCGACCCAGTCAACCAAGTATATCA 1861
Db 5196 ACAGACTGGGCGCTTTCAGAAATGAAGTCAAGCTGAGCGACCCAGTCAACCAATATATCA 5255
Qy 1862 TGACATGATGTCGGCTGACCTGAGGTCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1921
Db 5256 TGACATGATGTCGGCGGACCTGAGGTCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 5315
Qy 1922 TTCTGGCTGCTTGGCGCGATTTGCGCTATCCAGAGTGTGCTGCTATAGTAGTAGTAGGA 1981
Db 5316 TCCTGGCTGCTTGGCGCGATTTGCGCTATCCAGAGTGTGCTGCTATAGTAGTAGTAGGA 5375
Qy 1982 TTGTCTTGTCCGAAAGCCGCAATCATACCCGACAGGAGTCTCTTACCGGAGTTGCG 2041
Db 5376 TCGTCTTGTCCGGAAGCCGCAATCATACCTGACAGGAGTCTCTTACCGAGTTGCG 5435
Qy 2042 ATGAAATGGAAGTGTCT 2059
Db 5436 ATGAGATGGAAGTGTCT 5453

RESULT 15
AR166930
LOCUS AR166930 9379 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6284249.
ACCESSION AR166930
VERSION AR166930.1 GI:16243325
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 9379)
AUTHORS Barban, V.
TITLE Fusion polypeptide having the C protein and E1 protein of hepatitis C virus

JOURNAL Patent: US 6284249-A 1 04-SEP-2001;
FEATURES Location/Qualifiers
source 1. 9379
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 86.7%; Score 1786; DB 6; Length 9379;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
Qy 2 TGGCGCTATACGGCTATATGCCAGCAGACAAGGGCCCTTTGGGATGATATATACCA 61
Db 3396 TGGCGCCCATACGGCGTACGCCAGCAGACAAGGGCCCTCTAGGATGATATATACCA 3455
Qy 62 GCTTACCGCGCGGACAAAAACCCAGGTGAGGTTGAGTTCAGATCGTGTCAACTGCTG 121
Db 3456 GCCTACTGCGCGGACAAAAACCCAGGTGAGGTTGAGTTCAGATGTTGTCAACTGCTG 3515
Qy 122 CCAGACTTTCTTTGGCAACCTGATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCCG 181
Db 3516 CCAGAACCTTCTTGGCAACCTGATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCCG 3575
Qy 182 GAACAGGACCATGCGTCACTTAAGGTCCTGTTATCCAGATGTACACCAATGTGACC 241
Db 3576 GAACAGGACCATGCGTCACTTAAGGTCCTGTTATCCAGATGTATACCAATGTGACC 3635
Qy 242 AAGACCTCTGATGCTGCGCGCTCCCAAGGTGCGCTCATTAACACCATGCACTTTGCG 301
Db 3636 AAGACCTCTGATGCTGCGCGCTCCCAAGGTGCGCTCATTAACACCATGCACTTTGCG 3695
Qy 302 GCTCTCGGACCTTTACTCTGTCACGAGGACGCGCATGTCTCTTGTGCGCGGACGGG 361
Db 3696 GCTCTCGGACCTTTACTCTGTCACGAGGACGCGCATGTCTCTTGTGCGCGGACGGG 3755
Qy 362 GTATGCGAGGCGGAGCTGCTTTGCGCGCGCTATCTCTTACTTGAAGGCTCTCTCGG 421
Db 3756 GTATGCGAGGCGGAGCTGCTTGTGCGCGCGCTATCTCTTACTTGAAGGCTCTCTCGG 3815
Qy 422 GAGCGCTCTGCTGTCGCCCGCAGGACATGCGGTAGGACATATTACAGAGCCGCGTATGCA 481
Db 3816 GGGTCTCGCTGTTGTGCGCGCGGACGCGGTAGGACATATTAGGGCGCGCGTGTGCA 3875
Qy 482 CCGTGGAGTGGCTTAAGCGCGTGGACTTTCATCCCGGTAGAGAGCTTAGAGACCAACATGA 541
Db 3876 CCGTGGAGTGGCTTAAGCGCGTGGACTTTCATCCCTGTGAGAACCTTAGAGACCAACATGA 3935
Qy 542 GGTCCCGGCTGCTCAGACAACCTCTCCCAACAGAGTGTGCGCGAGAGCTACCAAGTGG 601
Db 3936 GGTCCCGGCTGCTCAGCGATACTCTCTCTCCAGAGTGTGCGCGAGAGCTTCCAGGTGG 3995
Qy 602 CCACCTGCTATGCTCCCGCGGCGGTAAAGACCAAGGTCCTCCCGCGCATACCGAG 661
Db 3996 CTACCTTCATGCTCCCGCGGCGGTAAAGACCAAGGTCCTCCCGCGCATACCGAG 4055
Qy 662 CTCAGGGCTACAAGGTCGTGCTCAACCCCTCGTGTGCTGCAACAAATGGGCTTTGGTG 721
Db 4056 CTCAGGGCTACAAGGTCGTGCTCAACCCCTCGTGTGCTGCAACAAATGGGCTTTGGTG 4115
Qy 722 CTTACATGTCAGGCGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
Db 4116 CTTACATGTCAGGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4175
Qy 782 CTACTGCGAGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 841
Db 4176 CCACTGCGAGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4235
Qy 842 CAGGGGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901
Db 4236 CAGGGGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4295
Qy 902 TCTTGGGCTATGCTGCTTCTTGAACAGAGAGACCGCGGGGCGAGACTGACTGTGC 961

```
Db 4296 TCTTGGGCACTGGGCACTGTCTCTTGAACCAAGCAGAGACTGGGGGGCAGAGCTGGTTGTGC 4355
Qy 962 TCGCCACCGCTACCCCTCCGGGCTCCGTCACGTGCCCCCATCTTAACATCGAGAGGTTG 1021
Db 4356 TCGCCACCGCACCCCTCCGGGCTCCGTCACGTGCCCCCATCCAACTCGAGAGGTTG 4415
Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGSCAAAGGCTATTCCCTTTGAAGCAATTA 1081
Db 4416 CTCTGTCCACCCCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTTCGAAGTAATCA 4475
Qy 1082 AGGGGGGAGACATCTCTCATCTTCTGCCACTCAAGAAGAGTGGCAGAGCTCGCGCAA 1141
Db 4476 AGGGGGGAGACATCTCTCATCTTCTGTCAATTCAAGAAGAGTGGCAGAACTCGCGCAA 4535
Qy 1142 AACTGGTTCGGGCTCAATGCGGTGGCTTACTACCGGCGCTTGATGTGTCGTCA 1201
Db 4536 AGCTGGTCGATTTGGGCATCAATGCCGTGGCTTACTACCGGCTTGGACGTGTCGTCA 4595
Qy 1202 TCCGACCAAGTGTGACGTTGTGCTGTGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261
Db 4596 TCCGACCAAGCGGATGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGCTATACCG 4655
Qy 1262 GCGACTTCGATTTCGCTGATAGACTGCAACAGGTGTGTACCCAGACAGTTCAGCTCAGCC 1321
Db 4656 GCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTTCAGCTCAGCC 4715
Qy 1322 TTGACCCCTACTCTCACCATTTGAGACAATCACGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381
Db 4716 TTGACCCCTACTCTCACCATTTGAGACAATCACGCTTCCCGAGGATGCTGTCTCCCGACTC 4775
Qy 1382 AACGTTCGGGTAGGACTTGGCAGAGGGAAGCAGGCACTTACAGATTTGTGGCACCGGGG 1441
Db 4776 AACGTTCGGGCAGGACTTGGCAGGGAAGCAGGCACTTACAGATTTGTGGCACCGGGG 4835
Qy 1442 AGCGTCTCTTCGCGATGTTTGACTGCTGTCTCTCGAGTGTGTATGACCGGGTGTG 1501
Db 4836 AGCGCCCTCCCGCATGTTCCGACTGCTGCTCGTCTCTGTGAGTGTGTATGACGAGGCTGTG 4895
Qy 1502 CTTGGTATGAGCTTACGCCCGCAGACACAGTTAGGCTACGAGCATACATGAACACCC 1561
Db 4896 CTTGGTATGAGCTACGCCCGCAGACACAGTTAGGCTACGAGCGTACATGAACACCC 4955
Qy 1562 CGGGACTTCCCGTGTGCCAAGACCATCTGAAATTTTGGGAGGCGTCTTTACGGGTCTCA 1621
Db 4956 CGGGGCTTCCCGTGTGCCAGGACCATCTGAAATTTTGGGAGGCGTCTTTTACAGGCTCA 5015
Qy 1622 CCCACATAGACGCCCACTTCTATCCAGACAAGCAGAGTGGGAAAACCTTCCCTATC 1681
Db 5016 CTATATAGATGCCCACTTCTATCCAGACAAGCAGAGTGGGGAACCTTCTTACC 5075
Qy 1682 TGGTAGCGTACCAAGCCAGCTGTGCGTAGAGCTCAAGCCCTCCCGCGTGGGGACC 1741
Db 5076 TGGTAGGCTACCAAGCCAGCTGTGCGTAGGCTCAAGCCCTCCCGCATCGTGGGACC 5135
Qy 1742 AGATGTGAAGTGTGTTGATCCGCTCAAGCCCAACCTTCCATGGGCCAACAACCTCTGCTAT 1801
Db 5136 AGATGTGAAGTGTGTTGATTCGCTCAAGCCCAACCTTCCATGGGCCAACAACCTCTGCTAT 5195
Qy 1802 ATAGACTGGGCGCTGTCCAGAAATGAGTCAACCTGACGCCACCCAGTCAACCAAGTATATCA 1861
Db 5196 ACAGACTGGGCGCTGTTCAGAAATGAAATCAACCTGACGCCACCCAGTCAACCAATATATCA 5255
Qy 1862 TGACATGATGTGCGGTGACTGGAGGTGCTACGAGTACTGGGTGCTGTTGGCGGCG 1921
Db 5256 TGACATGATGTGCGCGGACTGGAGGTGCTACGAGCACTGGGTGCTGTTGGCGGCG 5315
Qy 1922 TTCTGGCTGTTTGGCGCGGTATGCTATCCAGGCTGCGGTGCTATAGTAGGTAGGA 1981
Db 5316 TCCTGGCTGTTTGGCGCGGTATGCTATCCAGGCTGCGGTGCTATAGTAGGTAGGA 5375
Qy 1982 TTCTGTTGCCGGAAGCGGCAATCATACCCGACAGGAGTCTCTTACCGGAGTTGCG 2041
Db 5376 TCCTGTTGCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGGAGTTGCG 5435
```

Qy 2042 ATGAAATGGAAGAGTGCT 2059

Db 5436 ATGAGATGGAAGAGTGCT 5453

Search completed: February 27, 2005, 20:54:59

Job time : 8691 secs

The present invention relates to novel hepatitis C virus (HCV) NS3/4A proteins and their corresponding polynucleotides. NS3/4A sequences are useful for identifying the presence or absence of HCV in a subject. They are useful for preparing a medicament used for treating or preventing HCV infection. Sequences of the invention are also used as vaccines. The present sequence is a DNA encoding HCV NS3/4A protein			
XX	Sequence	2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;	
Query Match 100.0%; Score 2061; DB 6; Length 2061;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGCGCGCTATACGGCTATGCCAGCAGACAAAGGGCGCTTTGGATGCAATACACC	60
DB	1	ATGCGCGCTATACGGCTATGCCAGCAGACAAAGGGCGCTTTGGATGCAATACACC	60
QY	61	AGCTTGACCGCGCGGACAAAAACCAAGGTGAGGGTGAGATTGAGATCGTGTCAACTGCT	120
DB	61	AGCTTGACCGCGCGGACAAAAACCAAGGTGAGGGTGAGATTGAGATCGTGTCAACTGCT	120
QY	121	GCCAGACTTTCTTGGCAACTGCATTAACGGGGTGTGTGGACTGTCTACCATGAGCC	180
DB	121	GCCAGACTTTCTTGGCAACTGCATTAACGGGGTGTGTGGACTGTCTACCATGAGCC	180
QY	181	GGAAACAGGACCAATTCGTCAACCTTAAGGGTCTGTATTCCAGATGTACACCATGTGGAC	240
DB	181	GGAAACAGGACCAATTCGTCAACCTTAAGGGTCTGTATTCCAGATGTACACCATGTGGAC	240
QY	241	CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCGCGCTCAATTAACACCATGCACTTGC	300
DB	241	CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCGCGCTCAATTAACACCATGCACTTGC	300
QY	301	GGCTCTCGAACCTTTACCTGTACAGGACGCGCGATGTCAATTCCTGTGCGCGCACGG	360
DB	301	GGCTCTCGAACCTTTACCTGTACAGGACGCGCGATGTCAATTCCTGTGCGCGCACGG	360
QY	361	GGTGTAGGACGGGACGCTGCTTTCGCGCGCGCTATCTTCTTACCTTGAAGGCTCCTCG	420
DB	361	GGTGTAGGACGGGACGCTGCTTTCGCGCGCGCTATCTTCTTACCTTGAAGGCTCCTCG	420
QY	421	GGAGGCGCTCTGCTGTGCCCGCAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGC	480
DB	421	GGAGGCGCTCTGCTGTGCCCGCAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGC	480
QY	481	ACCGGTGAGTGCTTAAGGCGGTGGACTTCATCCCGTAGAGCTTAGAGACAAACCATG	540
DB	481	ACCGGTGAGTGCTTAAGGCGGTGGACTTCATCCCGTAGAGCTTAGAGACAAACCATG	540
QY	541	AGGTCCCGGTGTCTCAGACAACTCCTCCCGACAGCAGTGCCTCCAGAGCTACCAAGTG	600
DB	541	AGGTCCCGGTGTCTCAGACAACTCCTCCCGACAGCAGTGCCTCCAGAGCTACCAAGTG	600
QY	601	GCCACCTGCTATGCCCGCAGCGGTAAAGACCAACAGGTCCCGCGCATACGCA	660
DB	601	GCCACCTGCTATGCCCGCAGCGGTAAAGACCAACAGGTCCCGCGCATACGCA	660
QY	661	GCTCAGGGCTACAAGGTGCTGTGCTCAACCCCTCGTGTGCTGCAACATGGGCTTTGGT	720
DB	661	GCTCAGGGCTACAAGGTGCTGTGCTCAACCCCTCGTGTGCTGCAACATGGGCTTTGGT	720
QY	721	GCTTACATGTCCAGGCGCATGGATTGATCTTAAACATCAGGACTGGGGTGAGACAAT	780
DB	721	GCTTACATGTCCAGGCGCATGGATTGATCTTAAACATCAGGACTGGGGTGAGACAAT	780
QY	781	ACTACTGGCAGCCCGATCACTATTCACCTACGCAAGTTCCTTGGCGCGGGGTGT	840
DB	781	ACTACTGGCAGCCCGATCACTATTCACCTACGCAAGTTCCTTGGCGCGGGGTGT	840
QY	841	TCAGGGGGTCTTATGACATAAATTTGTGACAGGTGCCACTCCAGCGATGCAACATCC	900
DB	841	TCAGGGGGTCTTATGACATAAATTTGTGACAGGTGCCACTCCAGCGATGCAACATCC	900

QY	901	ATCTTTGGGCATTGGCACTGTCTCTTGACCAAGCAGAGACCGCGGGCGAGACTGTG	960
DB	901	ATCTTTGGGCATTGGCACTGTCTCTTGACCAAGCAGAGACCGCGGGCGAGACTGTG	960
QY	961	CTCGCACCGCTACCGCTCCGGGCTCCGTCACCTGTGCCCCATCCTTAACATCAGGAGGTT	1020
DB	961	CTCGCACCGCTACCGCTCCGGGCTCCGTCACCTGTGCCCCATCCTTAACATCAGGAGGTT	1020
QY	1021	GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTCTGAAGCAATT	1080
DB	1021	GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTCTGAAGCAATT	1080
QY	1081	AAGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGCTGCGCGCA	1140
DB	1081	AAGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGCTGCGCGCA	1140
QY	1141	AAACTGTGTCGGTGGGCGTCAATGCGGTGGCTTACTACCGGGCGCTTGTGTCCGTC	1200
DB	1141	AAACTGTGTCGGTGGGCGTCAATGCGGTGGCTTACTACCGGGCGCTTGTGTCCGTC	1200
QY	1201	ATCCCGACAGTGTGACGTTGTCGTGCGGCAACTGACGCGCTCATGACCGGCTTTTACC	1260
DB	1201	ATCCCGACAGTGTGACGTTGTCGTGCGGCAACTGACGCGCTCATGACCGGCTTTTACC	1260
QY	1261	GGCGACTTCGATTTCGGTGATAGACTGCAACACGCTGTGTCAACAGACAGTTCGATTCAGC	1320
DB	1261	GGCGACTTCGATTTCGGTGATAGACTGCAACACGCTGTGTCAACAGACAGTTCGATTCAGC	1320
QY	1321	CTTGACCTTACCTTACCATTTGAGACATTCAGCTTCCCGGATGCTGTCTCCCGTACT	1380
DB	1321	CTTGACCTTACCTTACCATTTGAGACATTCAGCTTCCCGGATGCTGTCTCCCGTACT	1380
QY	1381	CAACGTGCGGGTAGGACTGGCAGAGGAGCAGGCACTACAGATTTCGTGGCACCGGG	1440
DB	1381	CAACGTGCGGGTAGGACTGGCAGAGGAGCAGGCACTACAGATTTCGTGGCACCGGG	1440
QY	1441	GAGCGTCTTCTGGCACTGTTTGACTGCTGCTCTCTGCGAGTGTATGACGCGGGTTGT	1500
DB	1441	GAGCGTCTTCTGGCACTGTTTGACTGCTGCTCTCTGCGAGTGTATGACGCGGGTTGT	1500
QY	1501	GCTTGATATGAGCTTAGCGCCCGGAGACCAAGTTAGGCTACGACATACATGACACC	1560
DB	1501	GCTTGATATGAGCTTAGCGCCCGGAGACCAAGTTAGGCTACGACATACATGACACC	1560
QY	1561	CGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGCTCTTACGGGTCTC	1620
DB	1561	CGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGCTCTTACGGGTCTC	1620
QY	1621	ACCCACATAGACGCCCATTTCTTATCCAGACAAAGCAGAGTGGGAAAACCTTCCCTAT	1680
DB	1621	ACCCACATAGACGCCCATTTCTTATCCAGACAAAGCAGAGTGGGAAAACCTTCCCTAT	1680
QY	1681	CTGGTAGCTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGCTGGGAC	1740
DB	1681	CTGGTAGCTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGCTGGGAC	1740
QY	1741	CAGATGTGAAGTGTGTTGATCCGCTCAAGCCACCCCTCCATGGGCAACACCTCTGCTA	1800
DB	1741	CAGATGTGAAGTGTGTTGATCCGCTCAAGCCACCCCTCCATGGGCAACACCTCTGCTA	1800
QY	1801	TATAGACTGGGGCTGTCAGAAATGAAGTCACTGACGACCCAGTACCAGATATATC	1860
DB	1801	TATAGACTGGGGCTGTCAGAAATGAAGTCACTGACGACCCAGTACCAGATATATC	1860
QY	1861	ATGACATGATGTGCGCTGACCTGAGGTCGTCAAGAGTACCTGGGTGCTCGTGGCGG	1920
DB	1861	ATGACATGATGTGCGCTGACCTGAGGTCGTCAAGAGTACCTGGGTGCTCGTGGCGG	1920
QY	1921	GTTCTGGCTGCTTGGCGCGGATTTCCCTATCCAGAGGTGCGTGTCTATAGTAGGTAGG	1980
DB	1921	GTTCTGGCTGCTTGGCGCGGATTTCCCTATCCAGAGGTGCGTGTCTATAGTAGGTAGG	1980
QY	1981	ATTGTCTTGTCCGAAAGCGCGCAATCATATCCCGACAGGGGAAGTCTCTACCGGGAGTTC	2040

Db	1981	ATTGTCCTGTCGCGAAGCCGGCAATCATACCCAGACAGGGAAGTCCTCTACCGGAGTTC	2040			
Qy	2041	GATCAAAATGGAAGAGTGTCTGA	2061			
Db	2041	GATCAAAATGGAAGAGTGTCTGA	2061			
RESULT 2						
AAD31767						
ID	AAD31767 standard; DNA; 2061 BP.					
XX						
AC	AAD31767;					
XX						
XX	18-JUN-2002 (first entry)					
XX	Hepatitis C virus (HCV) NS3/4A DNA coding region.					
XX						
KW	Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;					
KW	cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;					
XX	ds.					
XX						
OS	Hepatitis C virus.					
XX						
FH	Key	Location/Qualifiers				
FT	CDS	1..2061				
FT		/*tag= a				
FT		/product= "HCV NS3/4A protein"				
XX						
PN	W0200213855-A2.					
XX						
PD	21-FEB-2002.					
XX						
PF	15-AUG-2001; 2001WO-IB001808.					
XX						
PR	17-AUG-2000; 2000US-0225767P.					
PR	29-AUG-2000; 2000US-0229175P.					
PR	03-NOV-2000; 2000US-00705547.					
XX						
PA	(TRIP-) TRIPEP AB.					
PI	Sallberg M, Hultgren C;					
XX						
DR	WPI; 2002-241837/29.					
XX						
DR	P-PSDB; AAE19900.					
XX						
PT	Vaccine compositions for treating and preventing disease, preferably					
PT	Hepatitis C virus infection, comprises ribavirin and antigen that has					
PT	epitope present in hepatitis C virus.					
XX						
PS	Claim 1; Page 94-95; 120pp; English.					
XX						
CC	The invention relates to a composition comprising ribavirin and an					
CC	antigen preferably non structural 3 protein (NS3)/4A fragment of					
CC	hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV					
CC	sequence. The composition is useful for enhancing an immune response to a					
CC	hepatitis C antigen in humans, domestic, sport or pet species and as					
CC	vaccines for treating and preventing HCV infections. The composition is					
CC	also useful for treating viral, bacterial, fungal diseases and cancer.					
CC	The present sequence is HCV NS3/4A DNA coding region					
XX						
SQ	Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;					
Query Match 100.0%; Score 2061; DB 6; Length 2061;						
Best Local Similarity 100.0%; Pred. No. 0;						
Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	1	ATGCGCGCTATCAGCGCTATGCCAGACAGAGGGGCCCTTTGGGATGCATAATCACC	60			
Db	1	ATGCGCGCTATCAGCGCTATGCCAGACAGAGGGGCCCTTTGGGATGCATAATCACC	60			
Qy	61	AGCTTGTACCGCGCGGGGCAAAAAACAGGTGGAGGGTTCAGATCGTGTCAACTGCT	120			

Qy 421 GGAGGCCCTCTGCTGTGTCGCCGAGGACANCCCGTAGGCATATTTCAGAGCCGCGGTATGC 480
Db 421 GGAGGCCCTCTGCTGTGTCGCCGAGGACATGCCGTAGGCATATTTCAGAGCCGCGGTATGC 480
Qy 481 ACCGCTGGAGTGGCTTAAGGCGGTGGACTTTCATCCCGTAGAGCTTAGAGCAACCATG 540
Db 481 ACCGCTGGAGTGGCTTAAGGCGGTGGACTTTCATCCCGTAGAGCTTAGAGCAACCATG 540
Qy 541 AGGTCCCGGTGTTCTTCAGACAACTCTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG 600
Db 541 AGGTCCCGGTGTTCTTCAGACAACTCTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG 600
Qy 601 GCCCACTGCAATGCTCCCAACGCGAGGGTAAAGACACCAAGGTCCGCGCGCATACGCA 660
Db 601 GCCCACTGCAATGCTCCCAACGCGAGGGTAAAGACACCAAGGTCCGCGCGCATACGCA 660
Qy 661 GCTCAGGCTACAGGTGCTGCTCAACCCCTCCGTTGCTGCAACATGGGCTTTGGT 720
Db 661 GCTCAGGCTACAGGTGCTGCTCAACCCCTCCGTTGCTGCAACATGGGCTTTGGT 720
Qy 721 GCTTACATGTCACAGGCCCATGGATTGATCCTTAACATCAGACTGGGGTGAGGACAATT 780
Db 721 GCTTACATGTCACAGGCCCATGGATTGATCCTTAACATCAGACTGGGGTGAGGACAATT 780
Qy 781 ACTACTGGCAGCCGATCACTGTTCCCACTTACCGGCAAGTTCCTTGGCGAGCGGGTGT 840
Db 781 ACTACTGGCAGCCGATCACTGTTCCCACTTACCGGCAAGTTCCTTGGCGAGCGGGTGT 840
Qy 841 TCAGGGGGTCTTATGACATAAATTTGTGACAGTGCCTCAACGATGCAACATCC 900
Db 841 TCAGGGGGTCTTATGACATAAATTTGTGACAGTGCCTCAACGATGCAACATCC 900
Qy 901 ATCTTTGGCACTGTCCTTGACCAAGCAGACCGGGGGGAGACTGACTGTG 960
Db 901 ATCTTTGGCACTGTCCTTGACCAAGCAGACCGGGGGGAGACTGACTGTG 960
Qy 961 CTCGCCACCGTACCCCTCGGGTCCGTCATCTGTGCCCACTCCTTAACATCGAGAGGTT 1020
Db 961 CTCGCCACCGTACCCCTCGGGTCCGTCATCTGTGCCCACTCCTTAACATCGAGAGGTT 1020
Qy 1021 GCTCTGTCCACTACGGAGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAATT 1080
Db 1021 GCTCTGTCCACTACGGAGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAATT 1080
Qy 1081 AAGGGGGGAGACATCTCATCTCTGCCACTCAAGAGAGTGCAGAGCTCGCGCA 1140
Db 1081 AAGGGGGGAGACATCTCATCTCTGCCACTCAAGAGAGTGCAGAGCTCGCGCA 1140
Qy 1141 AAACGTGTGCGTGGGCGTCAATGCGGTGGCTTACTACCGCGGCTTGTGTCGGTC 1200
Db 1141 AAACGTGTGCGTGGGCGTCAATGCGGTGGCTTACTACCGCGGCTTGTGTCGGTC 1200
Qy 1201 ATCCGACAGTGTGAGCTGTGTCGTGGCAATGACGCGCTCATGACGCGTTTACC 1260
Db 1201 ATCCGACAGTGTGAGCTGTGTCGTGGCAATGACGCGCTCATGACGCGTTTACC 1260
Qy 1261 GGCGACTTCGATTCGCTGATAGATGCAACACGTTGTACCCAGACAGTGCCTCAGC 1320
Db 1261 GGCGACTTCGATTCGCTGATAGATGCAACACGTTGTACCCAGACAGTGCCTCAGC 1320
Qy 1321 CTTGACCCCTACTTTCACCATTTAGACAAATCAGCTTCCCAAGGATGCTGTCCTCCGTACT 1380
Db 1321 CTTGACCCCTACTTTCACCATTTAGACAAATCAGCTTCCCAAGGATGCTGTCCTCCGTACT 1380
Qy 1381 CAACGTGGGGTAGGACTGGCAGAGGAGGACCAAGGATCTACAGATTGTGGCAACCGGG 1440
Db 1381 CAACGTGGGGTAGGACTGGCAGAGGAGGACCAAGGATCTACAGATTGTGGCAACCGGG 1440
Qy 1441 GAGCGTCTCTCTGCACTGTTGACTCGTCTGCTCTGCGAGTCTATGACGCGGGTGT 1500
Db 1441 GAGCGTCTCTCTGCACTGTTGACTCGTCTGCTCTGCGAGTCTATGACGCGGGTGT 1500

Qy 1501 GCTTGTATGAGTTACGCCCGCGAGACCAACAGTTAGGCTACGAGCATATGAACACC 1560
Db 1501 GCTTGTATGAGTTACGCCCGCGAGACCAACAGTTAGGCTACGAGCATATGAACACC 1560
Qy 1561 CCGGACTTCCCGTGTCCAGACCATCTTGAATTTTGGAGGGCGTCTTTACGGGTCTC 1620
Db 1561 CCGGACTTCCCGTGTCCAGACCATCTTGAATTTTGGAGGGCGTCTTTACGGGTCTC 1620
Qy 1621 ACCCACAATAGACGCCCACTTCTATCCAGACCAAGCAGAGTGGGGAACCTTTCCCTAT 1680
Db 1621 ACCCACAATAGACGCCCACTTCTATCCAGACCAAGCAGAGTGGGGAACCTTTCCCTAT 1680
Qy 1681 CTGCTAGCTACCAAGCCACCGTGTGCTGCTAGACTCAAGCCCTCCCGCGCTGTTGGAC 1740
Db 1681 CTGCTAGCTACCAAGCCACCGTGTGCTGCTAGACTCAAGCCCTCCCGCGCTGTTGGAC 1740
Qy 1741 CAGATGTGGAAGTGTCTTGTATCCGCTCAAGCCCACTTCCATGGGCAACACCTCTGCTA 1800
Db 1741 CAGATGTGGAAGTGTCTTGTATCCGCTCAAGCCCACTTCCATGGGCAACACCTCTGCTA 1800
Qy 1801 TATAGACTGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCCCAAGTATATC 1860
Db 1801 TATAGACTGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCCCAAGTATATC 1860
Qy 1861 ATGACATGTATGTGCGTGTGACCTGAGGTGTCTACAGTACCTGGGTGCTCTGGTGGCGC 1920
Db 1861 ATGACATGTATGTGCGTGTGACCTGAGGTGTCTACAGTACCTGGGTGCTCTGGTGGCGC 1920
Qy 1921 GTTCTGCTCTTTGGCGGTATTTGCTATCCACAGCTGCGTGTGCTATAGTAGGTAGG 1980
Db 1921 GTTCTGCTCTTTGGCGGTATTTGCTATCCACAGCTGCGTGTGCTATAGTAGGTAGG 1980
Qy 1981 ATTGTCTTGTCCGAAAGCGGCAATCATATCCGACAGGGAAGTCTCTTACCGGAGTTC 2040
Db 1981 ATTGTCTTGTCCGAAAGCGGCAATCATATCCGACAGGGAAGTCTCTTACCGGAGTTC 2040
Qy 2041 GATGAAATGGAAGTGTCTGA 2061
Db 2041 GATGAAATGGAAGTGTCTGA 2061

RESULT 4
ADG47658
ID ADG47658 standard; DNA; 2061 BP.
XX AC ADG47658;
XX AC ADG47658;
DT 11-MAR-2004 (first entry)
XX HCV NS3/4A domain DNA.
DE ds; gene; immunogen; hepatitis C virus; HCV infection; vaccine.
XX Hepatitis C virus.
XX Key Location/Qualifiers
XX CDS 1..2061
XX FT /*tag= a
XX FT /product= "NS3/4A domain"
XX US2003206919-A1.
XX 06-NOV-2003.
XX 26-NOV-2002; 2002US-00307047.
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 15-AUG-2001; 2001US-00929955.
XX 15-AUG-2001; 2001US-00930591.
XX (SALL/) SALLBERG M.
XX

PI Sallberg M;
XX WPI; 2004-051480/05.
DR P-PSDB; ADG47659.
XX
PT New purified or isolated nucleic acid useful for enhancing an immune
PT response to a hepatitis C antigen comprises specific nucleotide sequences
PT and the amino acid sequences.
XX
PS Example 1; SEQ ID NO 1; 83pp; English.
XX
CC The invention relates to a purified or isolated nucleic acid. The
CC peptides are useful as immunogens for the treatment and prevention of
CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.
XX
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
Query Match 100.0%; Score 2061; DB 12; Length 2061;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCGGCTATACGGGCTATGCCAGCAGACAAGGGGCTTTTGGGATGATAATCACC 60
Db 1 ATGCGGCTATACGGGCTATGCCAGCAGACAAGGGGCTTTTGGGATGATAATCACC 60
Qy 61 AGCTTGACCGCGCGGACAAACCAGGTGAGGGTGAGTTGAGTCTGTTCAACTGCT 120
Db 61 AGCTTGACCGCGCGGACAAACCAGGTGAGGGTGAGTTGAGTCTGTTCAACTGCT 120
Qy 121 GCCCAGACTTCTTGGCAACCTGCATTAACGGGTGTTGGAGTGTCTACCATGGAGCC 180
Db 121 GCCCAGACTTCTTGGCAACCTGCATTAACGGGTGTTGGAGTGTCTACCATGGAGCC 180
Qy 181 GGAAACAGGACCAATGCGTCACTAAGGGTCTGTTATCCAGATGACCAATGTGGAC 240
Db 181 GGAAACAGGACCAATGCGTCACTAAGGGTCTGTTATCCAGATGACCAATGTGGAC 240
Qy 241 CAAGACCTCTAGGCTGGCGGCTCCCAAGGTGCGCGCTCATTAACCATGCACTTGC 300
Db 241 CAAGACCTCTAGGCTGGCGGCTCCCAAGGTGCGCGCTCATTAACCATGCACTTGC 300
Qy 301 GGCTCTCTCGACCTTTACCTGGTCAAGGACACCGCATGTCATCTCTGCGCCGACGG 360
Db 301 GGCTCTCTCGACCTTTACCTGGTCAAGGACACCGCATGTCATCTCTGCGCCGACGG 360
Qy 361 GGTGATGGCAGGGGACGCTGCTTTCGCGCGGCTATCTTACTTGAAGGCTCCTCG 420
Db 361 GGTGATGGCAGGGGACGCTGCTTTCGCGCGGCTATCTTACTTGAAGGCTCCTCG 420
Qy 421 GGAGGCGCTCTGCTGTCGCGCAGACATGCGTAGGCTATTCAGAGCCGGGTATGC 480
Db 421 GGAGGCGCTCTGCTGTCGCGCAGACATGCGTAGGCTATTCAGAGCCGGGTATGC 480
Qy 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGCTTAGAGCAACCATG 540
Db 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGCTTAGAGCAACCATG 540
Qy 541 AGTCCCGGCTGTTCTCAGACAACTCTCCCAACAGAGTGCGCCAGAGCTACCAAGTG 600
Db 541 AGTCCCGGCTGTTCTCAGACAACTCTCCCAACAGAGTGCGCCAGAGCTACCAAGTG 600
Qy 601 GCCCACTGATGCTCCCAACCGGCGGTAGAGCAACAGGTCCCGCGCGATACGCA 660
Db 601 GCCCACTGATGCTCCCAACCGGCGGTAGAGCAACAGGTCCCGCGCGATACGCA 660
Qy 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGCTGCTGCTCAACAAATGGCTTTGGT 720
Db 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGCTGCTGCTGCTCAACAAATGGCTTTGGT 720
Qy 721 GCTTACATGTCNAGGCCCATGGGATTGATCTTAACATCAGGACTGGGGTGAGGCAATT 780

Db 721 GCTTACATGTCNAGGCCCATGGGATTGATCTTAACATCAGGACTGGGGTGAGGCAATT 780
Qy 781 ACTACTGGCAGCCCGATCATCGTATTCACCTACGCAAGTTCCTTCCGACGCGGTGT 840
Db 781 ACTACTGGCAGCCCGATCATCGTATTCACCTACGCAAGTTCCTTCCGACGCGGTGT 840
Qy 841 TCAGGGGGTGTCTTATGACATAAATTTGTGACGAGTGCATCTCCACGATGCAATCC 900
Db 841 TCAGGGGGTGTCTTATGACATAAATTTGTGACGAGTGCATCTCCACGATGCAATCC 900
Qy 901 ATCTTGGGCAATGGCAGTCTCTTACCAAGCAGACACCGGGGCGGAGACTGACTGTG 960
Db 901 ATCTTGGGCAATGGCAGTCTCTTACCAAGCAGACACCGGGGCGGAGACTGACTGTG 960
Qy 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCATCTGTCGCCATCTTAACATCGAGGAGTT 1020
Db 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCATCTGTCGCCATCTTAACATCGAGGAGTT 1020
Qy 1021 GCTCTGCTCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTCCCTGAGCAATT 1080
Db 1021 GCTCTGCTCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTCCCTGAGCAATT 1080
Qy 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAAGTGCAGCAGCTCGCGCA 1140
Db 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAAGTGCAGCAGCTCGCGCA 1140
Qy 1141 AAACCTGTCGCTTGGGCTCAATGCGGTGCTTACTACCGGGCTTGTATGTCCGTC 1200
Db 1141 AAACCTGTCGCTTGGGCTCAATGCGGTGCTTACTACCGGGCTTGTATGTCCGTC 1200
Qy 1201 ATCCGACCAAGTGTGAGCTGTCGTCGGCAACTGACGCCCTCATGACCGCTTACC 1260
Db 1201 ATCCGACCAAGTGTGAGCTGTCGTCGGCAACTGACGCCCTCATGACCGCTTACC 1260
Qy 1261 GCGCATCTCGATTGCGTGTGATAGACTGCAACAGTGTGTCAACAGACAGTCACTTACG 1320
Db 1261 GCGCATCTCGATTGCGTGTGATAGACTGCAACAGTGTGTCAACAGACAGTCACTTACG 1320
Qy 1321 CTTGACCTTACCTTACCATTTGAGACATCAAGTTCCTCCAGGATGCTGTCTCCGCTACT 1380
Db 1321 CTTGACCTTACCTTACCATTTGAGACATCAAGTTCCTCCAGGATGCTGTCTCCGCTACT 1380
Qy 1381 CAACGTGCGGGTAGGACTGGCAGAGGAGCAGAGCATCTACAGATTTGTGACACCGGG 1440
Db 1381 CAACGTGCGGGTAGGACTGGCAGAGGAGCAGAGCATCTACAGATTTGTGACACCGGG 1440
Qy 1441 GAGCGTCTCTTGGCATGTTGACTCGTGTCTCTGCGAGTGTATGACGCGGTGT 1500
Db 1441 GAGCGTCTCTTGGCATGTTGACTCGTGTCTCTGCGAGTGTATGACGCGGTGT 1500
Qy 1501 GCTTGGTATGAGCTTACGCGCGCGAGACCAAGTTCAGGCTACGAGCATACATGAACACC 1560
Db 1501 GCTTGGTATGAGCTTACGCGCGCGAGACCAAGTTCAGGCTACGAGCATACATGAACACC 1560
Qy 1561 CCGGACTTCCCGTGTGCAAGACCATCTTGAATTTGGAGGGCTCTTTACGGGTCTC 1620
Db 1561 CCGGACTTCCCGTGTGCAAGACCATCTTGAATTTGGAGGGCTCTTTACGGGTCTC 1620
Qy 1621 ACCCACAATAGACGCCATCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTAT 1680
Db 1621 ACCCACAATAGACGCCATCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTAT 1680
Qy 1681 CTGGTAGCGTACCAAGCCCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGAC 1740
Db 1681 CTGGTAGCGTACCAAGCCCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGAC 1740
Qy 1741 CAGATGTGGAAGTGTGATCGCTCAAGCCACCTCCATGGGCCAACCTCTGCTA 1800
Db 1741 CAGATGTGGAAGTGTGATCGCTCAAGCCACCTCCATGGGCCAACCTCTGCTA 1800
Qy 1801 TATAGACTGGGCGTGTCCGAATGAAGTCAACCTGACGCAACCCAGTCAACAGTATATC 1860
Db 1801 TATAGACTGGGCGTGTCCGAATGAAGTCAACCTGACGCAACCCAGTCAACAGTATATC 1860

Db 1801 TATAGACTGGGGCTGTCCAGAAATGAAGTACACCTGACGACCCAGTCACCAAGTATATC 1860
Qy 1861 ATGACATGATATGTCGGCTACCTCGAGTGCTCAGAGTACCTGGTGCTCGTTGGCGGC 1920
Db 1861 ATGACATGATATGTCGGCTACCTCGAGTGCTCAGAGTACCTGGTGCTCGTTGGCGGC 1920
Qy 1921 GTTCTGGCTGCTTTGGCGCGGTATGCTATCCACAGGCTGGTGCTCATAGTAGTAGG 1980
Db 1921 GTTCTGGCTGCTTTGGCGCGGTATGCTATCCACAGGCTGGTGCTCATAGTAGTAGG 1980
Qy 1981 ATTGCTGTGTCGGAAGCCGGCAATCATATCCGACAGGGAAGTCTCTACCGGGAGTTT 2040
Db 1981 ATTGCTGTGTCGGAAGCCGGCAATCATATCCGACAGGGAAGTCTCTACCGGGAGTTT 2040
Qy 2041 GATGAATGGAAGTGTCTGA 2061
Db 2041 GATGAATGGAAGTGTCTGA 2061

RESULT 5

AAF83669

ID AAF83669 standard; DNA; 6299 BP.

XX AC

XX AAF83669;

XX DT

XX 23-JUL-2001 (first entry)

XX DE

XX HCV NS34A ORF comprising pCMV-NS34A nucleic acid sequence.

XX KW

XX HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;

XX KW catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;

XX KW immunotherapy; NS34A; ds.

XX OS

XX Synthetic.

XX OS

XX Hepatitis C virus.

XX FH

XX Key Location/Qualifiers

XX CDS 1990..4050

XX FT /*tag= a

XX FT /*product= "NS34A"

XX XX

XX W0200138360-A2.

XX XX

XX 31-MAY-2001.

XX XX

XX 22-NOV-2000; 2000WO-US032326.

XX XX

XX 24-NOV-1999; 99US-0167502P.

XX XX

XX (CHIR) CHIRON CORP.

XX XX

XX Colt D, Medina-Selby A, Selby M, Houghton M;

XX PI

XX WPI; 2001-343948/36.

XX DR

XX P-PSDB; AAB62633.

XX XX

XX Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as

XX PT a vaccine against HCV, comprises a polypeptide having a mutation that

XX PT functionally disrupts the catalytic domain of NS3.

XX PT

XX Disclosure; Fig 9; 340pp; English.

XX PS

XX CC

XX The invention relates to an isolated mutant non-structural (NS) Hepatitis

XX CC C virus (HCV) polypeptide, comprising a polypeptide having a mutation in

XX CC the catalytic domain of NS3, where the mutation functionally disrupts the

XX CC catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a

XX CC and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide

XX CC CC and polynucleotide (preferably DNA or a plasmid) compositions can be used

XX CC in vaccines against HCV and as diagnostics. The antibodies raised against

XX CC these polypeptides can also be used as diagnostics, or for passive

XX CC immunotherapy. The antibodies are also useful for isolating and

XX CC identifying HCV particles. The present sequence represents the nucleic

XX CC acid sequence of the pCMV-NS34A comprising the ORF encoding a NS34A

CC polypeptide

XX SQ

SQ Sequence 6299 BP; 1449 A; 1713 C; 1637 G; 1500 T; 0 U; 0 Other;

XX

XX Query Match 86.7%; Score 1787; DB 4; Length 6299;

XX Best Local Similarity 91.7%; Pred. No. 0;

XX Matches 1889; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 1 ATGCGCCTTATCAGCGCTATGCCAGCAGACAAAGGGCCCTTTGGGATGATATACACC 60

Db 1990 ATGCGCGCCATCAGCGGTACGCCAGCAGACAAAGGGCCCTTCCCTAGGTGTCATAATCACC 2049

Qy 61 AGCTTGACCGCGCGGGAACAAACACAGGTGAGAGGTGAGTTGAGTCTGTCACCTGCT 120

Db 2050 AGCTTAACTGCGCGGGAACAAACACAGGTGAGAGGTGAGTTGAGTCTGTCACCTGCT 2109

Qy 121 GCCCAGAGCTTTCTTGGCAACTGCAATTAACGGGGTGTGGAGCTGTCTACCATGGAGCC 180

Db 2110 GCCCAACCTTCTTGGCAACGTGCATCAATGGGGTGTGGAGCTGTCTACCATGGAGCC 2169

Qy 181 GGAACAGGACCAATGCGTCACCTAAGGGTCTCTTTATCCAGATGTACCAATGTGGAC 240

Db 2170 GGAACAGGACCAATGCGTCACCTAAGGGTCTCTTTATCCAGATGTACCAATGTGGAC 2229

Qy 241 CAAGACCTCGTAGCTGGCGCGCTCCCAAGGTGCCCGCTCATTTAAACACCATGACATTGC 300

Db 2230 CAAGACCTCGTAGCTGGCGCGCTCCCAAGGTGCCCGCTCATTTAAACACCATGACATTGC 2289

Qy 301 GGCTCTCGGACCTTTACCTGTCAGAGGACGCGCATGTCATTCTGTCGCGCGGCGG 360

Db 2290 GGCTCTCGGACCTTTACCTGTCAGAGGACGCGCATGTCATTCTGTCGCGCGGCGG 2349

Qy 361 GGTGATGCGAGGGGACGCTGCTTTCCGCCCGGCTATCTTACTTTGAAAGGCTCTCTCG 420

Db 2350 GGTGATGCGAGGGGACGCTGCTTTCCGCCCGGCTATCTTACTTTGAAAGGCTCTCTCG 2409

Qy 421 GGAAGCCCTCTGTGTGCCCGGAGGACATGCCGTAGGACATATTCAGAGCGCGCGGTATGC 480

Db 2410 GGAAGCCCTCTGTGTGCCCGGAGGACATGCCGTAGGACATATTCAGAGCGCGCGGTATGC 2469

Qy 481 ACCCGTGGAGTGGCTAAGGGGTGGACTTCATCCCGTAGAGAGCTTAGAGACACCATG 540

Db 2470 ACCCGTGGAGTGGCTAAGGGGTGGACTTCATCCCGTAGAGAGCTTAGAGACACCATG 2529

Qy 541 AGGTCCCGGCTGTCTCAGACCACTCTCCGCCACAGCAGTCCGCCAGAGTACCAAGTG 600

Db 2530 AGGTCCCGGCTGTCTCAGACCACTCTCCGCCACAGCAGTCCGCCAGAGTACCAAGTG 2589

Qy 601 GCCCACCTGCATGCTCCACCGGAGGACACCAAGGTCCCGGCGGACATACGCA 660

Db 2590 GCTCACCTGCATGCTCCACCGGAGGACACCAAGGTCCCGGCGGACATACGCA 2649

Qy 661 GCTCAGGGCTACAGGTGCTGGTGCTCAACCCCTCCGTTGCTGCAACAATGGGGCTTGGT 720

Db 2650 GCTCAGGGCTACAGGTGCTGGTGCTCAACCCCTCCGTTGCTGCAACAATGGGGCTTGGT 2709

Qy 721 GCTTACATGTCCAAGGCCATGGATTGATCTTAAACATCAGGACTGGGGTGGAGACAATT 780

Db 2710 GCTTACATGTCCAAGGCCATGGATTGATCTTAAACATCAGGACTGGGGTGGAGACAATT 2769

Qy 781 ACTACTGGCAGCCCGATCAGCTATTCCACTAGCGCAAGTTCCTTGGCGGCGGGTGT 840

Db 2770 ACTACTGGCAGCCCGATCAGCTATTCCACTAGCGCAAGTTCCTTGGCGGCGGGTGT 2829

Qy 841 TCAGGGGGTCTTATGACATAAATTTGTGACAGGTGCCACTCCAGGATGCAACATCC 900

Db 2830 TCAGGGGGTCTTATGACATAAATTTGTGACAGGTGCCACTCCAGGATGCAACATCC 2889

Qy 901 ATCTTGGGCTATGCACTGCTTTGACCAAGCAGAGACCGGGGGGAGACTGACATGTG 960

Db 2890 ATCTTGGGCTATGCACTGCTTTGACCAAGCAGAGACTGCGGGGGGAGACTGACATGTG 2949

Qy 961 CTCGCCACCGCTACCCCTCCGGGGTCCGTCACCTGTGCCCCCATCTCTAAACATCGAGGAGGTT 1020

Db 2950 CTGCCACCCGCCACCCCTCGGGCTCGCTCACTGTGCCCATCCCAACATCGAGGAGGTT 3009
Qy 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATCCCTTTGAAGCAATT 1080
Db 3010 GCTCTGTCCACCCGAGAGATCCCTTTTATGGCAAGGCTATCCCTTTGAAGCAATT 3069
Qy 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTCCGACGAGCTCGCCGCA 1140
Db 3070 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTCCGACGAGCTCGCCGCA 3129
Qy 1141 AACTGTGCGGTTGGCGGTCAATGCGGTGCTTATACCGCGGCTTGAATGTGCGTC 1200
Db 3130 AAGTGTGCTCAATTTGGGCATCAATGCGGTGCTTATACCGCGGTCTTGACGTGCGTC 3189
Qy 1201 ATCCGACCAAGTGGTACGTTGCTGTCGCAACTGACGCCCTCATGACCGGCTTTACC 1260
Db 3190 ATCCGACCAAGTGGTACGTTGCTGTCGCAACTGACGCCCTCATGACCGGCTTTACC 3249
Qy 1261 GGGGACTTCGATTCGGTGTAGACTGCAACACGCTGTGTCAACCAGACAGTCGACTTCAGC 1320
Db 3250 GGGGACTTCGATTCGGTGTAGACTGCAACACGCTGTGTCAACCAGACAGTCGACTTCAGC 3309
Qy 1321 CTTGACCCCTACCTTACCATGATGAGCAATCAACGCTTCCCGAGATGCTGTCTCCGTA 1380
Db 3310 CTTGACCCCTACCTTACCATGATGAGCAATCAACGCTTCCCGAGATGCTGTCTCCGTA 3369
Qy 1381 CAACGTCGGGTAGGCTGCGAGGAGGACCGAGGATCTACAGATTTGTGGACCGGG 1440
Db 3370 CAACGTCGGGTAGGCTGCGAGGAGGACCGAGGATCTACAGATTTGTGGACCGGG 3429
Qy 1441 GAGCGTCTTCTGGCATGTTGACTGCTGTCTGTCTGTCAGTGTATGACGGGGTGT 1500
Db 3430 GAGCGTCTTCTGGCATGTTGACTGCTGTCTGTCTGTCAGTGTATGACGGGGTGT 3489
Qy 1501 GCTTGTATGAGCTTACGCGCCCGGAGACACAGTGTAGGCTACGAGATACATGAAAC 1560
Db 3490 GCTTGTATGAGCTTACGCGCCCGGAGACTACAGTGTAGGCTACGAGATACATGAAAC 3549
Qy 1561 CCGGGACTTCCGCTGTGCCAAGACCATCTTGAATTTGGAGGGCGTCTTACGGGTCTC 1620
Db 3550 CCGGGACTTCCGCTGTGCCAAGACCATCTTGAATTTGGAGGGCGTCTTACGGGGTCTC 3609
Qy 1621 ACCCATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
Db 3610 ACTCATATAGATGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 3669
Qy 1681 CTGGTAGCTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCCCGGTGCTGGGAC 1740
Db 3670 CTGGTAGCTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCCCGGTGCTGGGAC 3729
Qy 1741 CAGATGTGAAGTGTGATCCCTCTCAAGCCCAACCTCCATGGGCCACACCTCTGCTA 1800
Db 3730 CAGATGTGAAGTGTGATCCCTCTCAAGCCCAACCTCCATGGGCCACACCTCTGCTA 3789
Qy 1801 TATAGACTGGCGCTGTCCAGATGAAGTCAACCTGACGACCCAGTCAACCAAGTATATC 1860
Db 3790 TACAGACTGGCGCTGTTCAGATGAATCAACCTGACGACCCAGTCAACCAAGTATATC 3849
Qy 1861 ATGACATGTATGTGCGGTGACCTGAGGTGCTCAAGAGTACCTGGGTGCTGGTGGCGC 1920
Db 3850 ATGACATGTATGTGCGGTGACCTGAGGTGCTCAAGAGTACCTGGGTGCTGGTGGCGC 3909
Qy 1921 GTTCTGCTGCTTTGGCGGTATTCCTTATCCAGAGTGTGCTGCTATAGTAGTAG 1980
Db 3910 GTTCTGCTGCTTTGGCGGTATTCCTTATCCAGAGTGTGCTGCTATAGTAGTAG 3969
Qy 1981 ATTGTCTTGTCCGGAAGCGCGCAATCATACCCGACAGGGAAGTCTCTACCGGGATTC 2040
Db 3970 GTGCTCTTGTCCGGAAGCGCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGATTC 4029
Qy 2041 GATGAATGGAAGAGTGCT 2059

Db 4030 GATGAGATGGAAGAGTGCT 4048
RESULT 6
AAN92106
ID AAN92106 standard; DNA; 7310 BP.
XX
AC AAN92106;
XX
DT 09-SEP-2004 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX
DE Combined ORFs of the HCV cDNAs from clones K9-1 through 15e.
XX Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH; ss.
XX Hepatitis C virus.
OS Unidentified.
XX Key Location/Qualifiers
FT CDS 3..7310
FT /*tag= a
XX
PN EP318216-A.
XX 31-MAY-1989.
PD
XX 18-NOV-1988; 88EP-00310922.
PF
XX 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 06-MAY-1988; 88US-00191263.
PR 26-OCT-1988; 88US-00263584.
PR 14-NOV-1988; 88US-00271450.
XX
(CHIR) CHIRON CORP.
PA (CHIR) CHIRON CORP.
PA Houghton M, Choo QL, Kuo G;
XX
PI WPI; 1989-159274/22.
DR P-P8DB; AAP92050.
XX
PT Purified hepatitis C virus - and associated nucleic acids and
PT polypeptide(s).
XX
PS Claim 3; Fig 47-1 - 47-8; 139pp; English.
XX
CC It is a double-stranded nucleotide sequence of the open reading frame
CC (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus
CC (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to
CC detect the presence of HCV nucleic acids in samples. The polypeptide(s)
CC it encodes could be used as immunoassay reagents and vaccines and to
CC generate antibodies useful in diagnosis and passive immunotherapy for HCV
CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR
CC field.) (Updated on 25-MAR-2003 to correct PI field.)
CC
CC Revised record issued on 09-SEP-2004 : Correction to keywords
XX
XX Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 0 U; 4 Other;
SQ
Query Match 86.7%; Score 1786; DB 1; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
Qy 2 TGGCGCCCTATCACGGCTATGCCAGACAGACAAAGGGGCCCTTTGGGATGCATAATCACCA 61
Db 1729 TGGCGCCCTATCACGGCTATGCCAGACAGACAAAGGGGCCCTCTAGGGTGCAATACCA 1788
Qy 62 GCTTGACCGCGCGGACAAAACCAAGGTGGAGGTTCAGATCGTGTCATCTGCTG 121

Db 1789 GCCTAACTGGCGGGGACAAAACCAAGTGGAGGGTGAGGTCAGANTGTGTCAACTGCTG 1848
Qy 122 CCCAGACTTTCTTGGCAACCTGCAATTAACGGGGGTGTGTGGACTGTGTACCATGGAGCG 181
Db 1849 CCCAAACCTTCTTGGCAACCTGCAATTAACGGGGGTGTGTGGACTGTGTACCATGGAGCG 1908
Qy 182 GAACAGAGACCATTTGCGTCACTTAAGGGTCTGTATCCAGATGTACACCAATGTGACC 241
Db 1909 GAACAGAGACCATTTGCGTCACTTAAGGGTCTGTATCCAGATGTATACCAATGTGACC 1968
Qy 242 AAGACCTCGTAGGCTGGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGGCG 301
Db 1969 AAGACCTTGTGGCTGGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGGCG 2028
Qy 302 GCTCCTCGGACCTTTACCTGTGACAGGACGCGCATGTCACTTCTGTGGCGGACGGG 361
Db 2029 GCTCCTCGGACCTTTACCTGTGACAGGACGCGCATGTCACTTCTGTGGCGGACGGG 2088
Qy 362 GTGATGGAGGGGACGCTGTCTTGGCGCGCGGCTATCTTACTTTGAAGGCTCTCGG 421
Db 2089 GTGATAGAGGGGACGCTGTCTTGGCGCGCGGCTATCTTACTTTGAAGGCTCTCGG 2148
Qy 422 GAGGCTCTGCTGTGCCCGCAGACATGCGTAGGCATATTGAGGCGCGGCTGTGCA 481
Db 2149 GGGTCCGCTGTGTGGCGCGGCGACGCGGTGGGCTATTTAGGGCGCGGCTGTGCA 2208
Qy 482 CCCGTGAGTGGCTAAAGCGGTGGACTTTCATCCCGGTAGAGAGCTTGAAGCAACCATGA 541
Db 2209 CCCGTGAGTGGCTAAAGCGGTGGACTTTCATCCCGGTAGAGAGCTTGAAGCAACCATGA 2268
Qy 542 GGTCCCGGTGTCTCAGACAACTCTTCCCAACAGAGTGTCCCGAGAGCTTACCAAGTGG 601
Db 2269 GGTCCCGGTGTCTCAGAGTAACCTCTTCCCAACAGAGTGTCCCGAGAGCTTCCAGTGG 2328
Qy 602 CCCACCTGCATGCTCCCAACGGCGGTAGAGACCAAGGTCGCGGCCCATACGAG 661
Db 2329 CTCACCTCCATGCTCCCAACGGCGGTAGAGACCAAGGTCGCGGCCCATACGAG 2388
Qy 662 CTCAGGGGTCAAGGTGTGTGTCTCAACCGCTCGTGTGTCACCAATGGGCTTTGGTG 721
Db 2389 CTCAGGGGTCAAGGTGTGTGTCTCAACCGCTCGTGTGTCACCAATGGGCTTTGGTG 2448
Qy 722 CTTACATGTCAAGGCCCATCGGATTTGATCTTAACATCAGAGCTGGGGTAGGCAATTA 781
Db 2449 CTTACATGTCAAGGCTCATGGATCGATCTTAACATCAGAGCTGGGGTAGGCAATTA 2508
Qy 782 CTACTGGGCGCGCATCAGCTATTCCACCTAGCGCAAGTTCCTTGGCGAGCGGGGTGT 841
Db 2509 CCACTGGGCGCGCATCAGCTATTCCACCTAGCGCAAGTTCCTTGGCGAGCGGGGTGT 2568
Qy 842 CAGGGGTGTCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCAACATCCA 901
Db 2569 CGGGGGGCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCCACATCCA 2628
Qy 902 TCTTGGGCAATTTGCACTGTCTTGAACAGAGAGCGCGGGGGGAGACTGACTGTC 961
Db 2629 TCTTGGGCAATTTGCACTGTCTTGAACAGAGAGCTGCGGGGGGAGACTGCTGTGTC 2688
Qy 962 TCGCCACCGCTACCGCTCCGGGCTCCGCTACCTGTCGCCCATCTTAACATCGAGAGGTTG 1021
Db 2689 TCGCCACCGCTACCGCTCCGGGCTCCGCTACCTGTCGCCCATCTTAACATCGAGAGGTTG 2748
Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTTCCCTTGAAGCAATTA 1081
Db 2749 CTCTGTCCACACCGGAGAGATCCCTTTTATGCAAGGCTATTTCCCTTGAAGCAATTA 2808
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGTGCAGAGCTCCCGCAA 1141
Db 2809 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGTGCAGAGCTCCCGCAA 2868
Qy 1142 AACTGTGCGGCTTGGGCGTCAATGCGGTGCTTACTACCGGGGCTTGAATGTGCGTCA 1201
Db 2869 AGCTGGTGGCATTTGGGCTCAATGCGGTGCTTACTACCGGGTCTTGAAGCTGTCGCTCA 2928

Qy 1202 TCCGACCAAGTGTGACGTTGTCTGTGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261
Db 2929 TCCGACCAAGTGTGACGTTGTCTGTGGCAACTGACGCCCTCATGACCGGCTTTACCG 2988
Qy 1262 GCGACTTGCATTCGGTGTATAGACTGCAACAGTGTGTACCCAGACAGTGCATTCAGCC 1321
Db 2989 GCGACTTGCATTCGGTGTATAGACTGCAATACGTTGTGTACCCAGACAGTGCATTCAGCC 3048
Qy 1322 TTGACCTTACCTTCCCAATTGAGACATCAAGCTTCCCGAGATGTGTCTCCCGTACTC 1381
Db 3049 TTGACCTTACCTTCCCAATTGAGACATCAAGCTTCCCGAGATGTGTCTCCCGTACTC 3108
Qy 1382 AACGTCGGGTGAGACTTGGCAGAGGGAAGCAGGCATCTACAGATTTGTGTGCAACCGGGG 1441
Db 3109 AACGTCGGGTGAGACTTGGCAGAGGGAAGCAGGCATCTACAGATTTGTGTGCAACCGGGG 3168
Qy 1442 AGCGTCTTCTGGGATGTTTGAAGCTGTCTCTCGAGTGTATGACGCGGGTGTG 1501
Db 3169 AGCGTCTTCTGGGATGTTTGAAGCTGTCTCTCGAGTGTATGACGCGGGTGTG 3228
Qy 1502 CTTGTATGAGCTTACCGCCCGCAGACCAAGTGTGCTACGAGCATACATGAACACCC 1561
Db 3229 CTTGTATGAGCTTACCGCCCGCAGACCAAGTGTGCTACGAGCGTACATGAACACCC 3288
Qy 1562 CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCTCTTTACGGGTCTCA 1621
Db 3289 CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCTCTTTACAGGCTCA 3348
Qy 1622 CCACATAGAGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 1681
Db 3349 CTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC 3408
Qy 1682 TGGTAGGTACCAAGCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
Db 3409 TGGTAGGTACCAAGCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 3468
Qy 1742 AGATGTGGAAGTGTGTATCCGTCTCAAGCCCAACCTTCCATGGGCCCAACCTCTGTAT 1801
Db 3469 AGATGTGGAAGTGTGTATCCGTCTCAAGCCCAACCTTCCATGGGCCCAACCTCTGTAT 3528
Qy 1802 ATAGACTGGGGCTGTCCAGATGCAAGTCACTTACGACGACCAAGTCAACCAAGTATATCA 1861
Db 3529 ACAGACTGGGGCTGTTCAGAAATGAAATCACTTACGACGACCAAGTCAACCAAGTATATCA 3588
Qy 1862 TGACATGATGTGCGCTGACCTGGAGTCTGACAGTACCTGGTGTCTGCTGCGGCG 1921
Db 3589 TGACATGATGTGCGCTGACCTGGAGTCTGACAGACCTGGGTGTCTGCTGCGGCG 3648
Qy 1922 TTCTGGTGTCTTTGGCGCGCTATTGCTTATCCAGGCTTGGTGTCTAGTAGGTAGGA 1981
Db 3649 TTCTGGTGTCTTTGGCGCGCTATTGCTTATCCAGGCTTGGTGTCTAGTAGGTAGGA 3708
Qy 1982 TTGTCTTGTCCGGAAGCCCGCAATCATACCCGACAGGAAGTCTCTTACCGGAGTTCG 2041
Db 3709 TCGTCTTGTCCGGAAGCCCGCAATCATACCTGACAGGAAGTCTCTTACCGAGATTCG 3768
Qy 2042 ATGAAATGGAAGAGTGTCT 2059
Db 3769 ATGAGATGGAAGAGTGTCT 3786

RESULT 7

AAN90336
ID AAN90336 standard; DNA; 7310 BP.
XX AAN90336;
XX AC
XX AC
!DT 25-MAR-2003 (revised)
DT 19-JUL-2001 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Composite hepatitis C virus (HCV) cDNA.

XX	Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.
KW	Pan troglodytes.
XX	GB2212511-A.
XX	26-JUL-1989.
XX	18-NOV-1988; 88GB-00027024.
XX	18-NOV-1987; 87US-00122714.
PR	30-DEC-1987; 87US-00139886.
PR	26-FEB-1988; 88US-00161072.
PR	26-OCT-1988; 88US-00263584.
XX	(CHIR) CHIRON CORP.
XX	Houghton M, Choo QL, Kuo G;
XX	WPI; 1989-215054/30.
DR	P-PSDB; AAP90288.
XX	Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT	polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT	infection.
XX	Disclosure; Fig 47; 30pp; English.
XX	The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by
CC	aligning clones k9-1 through 15e in 5'-3' direction. The cDNA encodes
CC	antigens which react with antibodies in patients with non-A non-B
CC	hepatitis (NANBH). The cDNA can be used to design probes, or to
CC	synthesize polypeptides, which are used to diagnose HCV-induced NANBH, to
CC	raise antibodies for immunoassay or treatment, or to produce vaccines.
CC	See also AAP90288, and AAN90303-35. (N.B. This record was resubmitted to
CC	correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR
CC	field.)
XX	Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 U; 0 Other;
SQ	
Query Match	86.7%; Score 1786; DB 1; Length 7310;
Best Local Similarity	91.7%; Pred. No. 0;
Matches 1888; Conservative	0; Mismatches 170; Indels 0; Gaps 0;
Qy	2 TGGCGCCTATCACGCGCTATGCCCGACAGACAGGGGCGCTTTGGGATGCATAATCACCA 61
Db	1729 TGGCGCCCATCACGCGGTACGCCCAGCAGACAGGGGCGCTCTTAGGGTGCAATAATCACCA 1788
Qy	62 GCTTGCACCGCGCGGACAAAACACAGGTGGAGGCTGAGGTTTCAGATCGTGTCAACTGCTG 121
Db	1789 GCCTAACTGGCGGACAAAACCAAGTGAGGCTGAGGTCAGATGTGTCACTGCTG 1848
Qy	122 CCCAGACTTCTTGGGCACTGTGCATTAACCGGGGTGTGTGGACTGTCTACCATGGAGCCG 181
Db	1849 CCCAAAGCTTCTGGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCAAGGGGCG 1908
Qy	182 GAACAAGACCATTCGCTCACTAAGGCTCTGTATCCAGATGTACCAATGTGAC 241
Db	1909 GAACGAGGACCATTCGCTGACCCAAAGGCTCTGTATCCAGATGTATACCAATGTGAC 1968
Qy	242 AAGACCTCTGTAGGCTGGCGCGCTCCCCAAGGTGCGCGCTCATTAACACCATGCACTTGG 301
Db	1969 AAGACCTTGTGGCTGGCGCGCTCCGCAAGGTAGCCGCTCATTGACACCTGCACTTGG 2028
Qy	302 GCTCTCGGACCTTTACTGTGTACAGGACGCGCGATGCAATTCCTGTGCGCGACGCG 361
Db	2029 GCTCTCGGACCTTTACTGTGTACAGGACGCGCGATGCAATTCCTGTGCGCGACGCG 2088
Qy	362 GTGATGCGAGGGGAGCGCTGTTTCGCGCGCGCTATCTCTTACTTGAAGGCTCTCTCG 421
Db	2089 GTGATGCGAGGGGAGCGCTGCTGTTCGCGCGCGCCCAATTCCTACTTGAAGGCTCTCTCG 2148

|||||
Db 3229 CTTGTATGAGCTACGCCCGCGGAGACTACAGTTAGGCTACGAGCTACATGACACCC 3288
1562 CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTGGAGGGCTCTTTACGGGTCTCA 1621
3289 CGGGCTTCCGTGTGCCAGGACCATCTTGAATTTGGAGGGCTCTTTACAGGCTCA 3348
1622 CCCATAGAGCCGCTCTCTATCCAGCAAAAGCAGAGTGGGAAACCTTCCCTATC 1681
3349 CTCATATAGATGCCACCTTCTATCCAGCAAAAGCAGAGTGGGAGAACCTTCTTACC 3408
1682 TGGTAGCTTACACAGCCAGCTGTGCGCTAGAGCTCAAGCCCTCCCGCTGTGGAC 1741
3409 TGGTAGCTTACCAAGCCAGCTGTGCGCTAGAGCTCAAGCCCTCCCGCTGTGGAC 3468
1742 AGATGTGGAAGTCTTGTATCCGCTCTCAAGCCACCCCTCCATGGGCCAACACTCTGCTAT 1801
3469 AGATGTGGAAGTCTTGTATCCGCTCAAGCCACCCCTCCATGGGCCAACACTCTGCTAT 3528
1802 ATAGACTGGGGCTGTCCAGAAATGAAGTCAACCTGACGCAACCAAGTCAACCAATATCA 1861
3529 ACAGACTGGGGCTGTCTCAGAAATGAAGTCAACCTGACGCAACCAAGTCAACCAATATCA 3588
1862 TGACATGTATCTCGGCTACCTGGAGTCTGTCAGAGTACCTGGTCTCTTGGCGCG 1921
3589 TGACATGTATCTCGGCTACCTGGAGTCTGTCAGAGTACCTGGTCTCTTGGCGCG 3648
1922 TTCTGGCTGTCTGGCCGCTATTGCCCTATCCAGGCTGCTGTATAGTAGTAGGA 1981
3649 TCCTGGCTGTCTGGCCGCTATTGCCCTATCCAGGCTGCTGTATAGTAGTAGGA 3708
1982 TTGTCTTGTCCGAAAGCCGCAATCATACCCGACAGGAAAGTCTCTTACCGGAGTTG 2041
3709 TCGTCTTGTCCGAAAGCCGCAATCATACCCGACAGGAAAGTCTCTTACCGAGATTG 3768
2042 ATGAATGGAGAGTGCT 2059
3769 ATGAGATGGAGAGTGCT 3786

RESULT 8

AAQ98221 standard; cDNA to mRNA; 7310 BP.

XX AC AAQ98221;
XX AC
DT 25-MAR-2003 (revised)
DT 15-AUG-1996 (first entry)
XX DE
XX DE Hepatitis C virus clone genome.
XX KW Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;
XX KW antibodies; immunoprophylaxis; sera; serum; ds.
XX OS

Hepatitis C virus.

US5443965-A.

22-AUG-1995.

05-APR-1991; 91US-00681703.

06-APR-1990; 90US-00505611.

09-OCT-1990; 90US-00594854.

(GENE-) GENELABS INC.

Moeckli R, Reyes GR, Kim JP;

WPI; 1995-302120/39.

XX New nucleic acids encoding hepatitis C virus antigens - used to develop
PT prods. for detection of HCV-infected sera and prodn. of vaccines and anti
PT

PT -HCV antibodies.

Example 4; Fig 11; 71bp; English.

CC Hepatitis C virus (HCV) antigens can be used for detecting HCV infected
CC sera and individuals infected with HCV. They can also be used in an anti-
CC HCV vaccine or for the production of anti-HCV antibodies which can be
CC used for passive immunoprophylaxis. The antigens consistently identify
CC more HCV positive serum samples with a high degree of specificity. See
CC AAQ98202-14 and AAQ98193-51. (Updated on 25-MAR-2003 to correct PF
CC field.) (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 7310 BP; 1494 A; 2217 C; 2060 G; 1539 T; 0 U; 0 Other;

Query Match 86.7%; Score 1786; DB 2; Length 7310;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGGCGCTATCAGCGCTATGCCAGCAGACAAGGGCGCTTTGGGATGCATATACCA 61

Db 1729 TGGCGCCCATCACGGCGTACGCCAGCAGACAAGGGCGCTCTAGGGTGCATAATCACA 1788

Qy 62 GCTTGACGGCGCGGACAAAACCAAGGTGGAGGTGAGTTCAGATCGTGTCAACTGCTG 121

Db 1789 GCCTAACTGGCGCGGACAAAACCAAGTGGAGGTGAGTTCAGATCGTGTCAACTGCTG 1848

Qy 122 CCCAGACTTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCGG 181

Db 1849 CCCAAAACCTTCTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCGG 1908

Qy 182 GAACAAGGACCATGTGCTACCTAAGGGTCTGTATTATCCAGATGTACACCAATGTGGACC 241

Db 1909 GAACGAGGACCATGTGCTACCTAAGGGTCTGTATTATCCAGATGTACACCAATGTGGACC 1968

Qy 242 AAGACCTCTGAGTGGCGCTCCCAAGGTGCGCTCATTAACACCATGCACTTGGC 301

Db 1969 AAGACCTCTGAGTGGCGCTCCCAAGGTGCGCTCATTAACACCATGCACTTGGC 2028

Qy 302 GCTCTCGGACCTTTACTGTCACAGGACGCGGATGTCTATTCCTGTGCGCGACGGG 361

Db 2029 GCTCTCGGACCTTTACTGTCACAGGACGCGGATGTCTATTCCTGTGCGCGACGGG 2088

Qy 362 GTGATGGCAGGGCAGCTGCTTTGCGCCCGGCTATCTCTTACTTGAAAGGCTCTCTCG 421

Db 2089 GTGATGGCAGGGCAGCTGCTTTGCGCCCGGCTATCTCTTACTTGAAAGGCTCTCTCG 2148

Qy 422 GAGCCCTCTGCTGTGCGCGCAGGACATGCGGTAGGCATATTACAGACCCCGGTATGCA 481

Db 2149 GGGTCCGCTGTTGTGCGCGGCGGACGCGGTGGGCATATTTAGGGCCGCGGTGTGCA 2208

Qy 482 CCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGGTAGAGAGCTTAGAGACAACCATGA 541

Db 2209 CCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGGTAGAGAGCTTAGAGACAACCATGA 2268

Qy 542 GGTCCCGGTGTTCTCAGACAACCTCTCCACACGAGTGGCCCCAGAGCTTACCAAGTGG 601

Db 2269 GGTCCCGGTGTTCAACGATAACTCTCTCCACAGTAGTGGCCCCAGAGCTTCCAGGTGG 2328

Qy 602 CCACCTGCTATGCTCCACCGCGGAGCGGTAAAGCACAAGGTCCCGCCGCATACGAG 661

Db 2329 CTCACCTCCATGCTCCACAGGCGGCAAAAGGTCACCAAGGTCCCGCTCATATGCA 2388

Qy 662 CTCAGGGCTACAAGTCTGTGCTCAACCCCTCGTGTGCTGCAACCAATGGGCTTTGGTG 721

Db 2389 CTCAGGGCTACAAGTCTGTGCTCAACCCCTCGTGTGCTGCAACCAATGGGCTTTGGTG 2448

Qy 722 -CTTACATGTCGAAGGCCCATGGGATGATCCTTAACATCAGGACTGGGGTGAAGCAATTA 781

Db 2449 CTTACATGTCGAAGGCCCATGGGATGATCCTTAACATCAGGACTGGGGTGAAGCAATTA 2508

Qy 782 CTACTGGCAGCCGATCAGGTATTCACCTACGGCAAGTCTCTTGGCCGACGCGGGTGT 841

Db 2509 CCACTGGCAGCCGATCAGGTATTCACCTACGGCAAGTCTCTTGGCCGACGCGGGTGT 2568

```
QY 842 CAGGGGTCCTTATGACATAATAATTGTGACGAGTGCCACTCCACGGATGCAACATCCA 901
Db |||||
QY 2569 CGGGGGCGCTTAAGACATAATAATTGTGACGAGTGCCACTCCACGGATGCAACATCCA 2628
Db |||||
QY 902 TCTTGGGCATTTGGCACTGTCTTTGACCAAGCAGAGACCGGGGGGCGAGACTGACTGTGC 961
Db |||||
QY 2629 TCTTGGGCATCGGCATGTCTTTGACCAAGCAGAGACTCGGGGGGCGAGACTGTTTGTGC 2688
QY 962 TCGCCACCGGTACCCCTCGGGGTCGTCACCTGTGCCCCATCTTAACATCGAGGAGTTG 1021
Db |||||
QY 2689 TCGCCACCGGCACCCCTCGGGGTCGTCACCTGTGCCCCATCTTAACATCGAGGAGTTG 2748
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGSCAAGGCTATTCCCTTTGAAGCAATTA 1081
Db |||||
QY 2749 CTCTGTCCACCACCGGAGAGATCCCTTTTATGSCAAGGCTATTCCCTTCGAAGTAATCA 2808
QY 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTGCGACGAGCTGCGCGCAA 1141
Db |||||
QY 2809 AGGGGGGAGACATCTCATCTTCTGTCAATCAAAGAAGAGTGCGACGAACTCGCGCAA 2868
QY 1142 AACTGTGCGGTGGGGTCAATGCCGTGGCTTACTACCGCGCCTTGATGTGTCCGTCA 1201
Db |||||
QY 2869 AGCTGTGCAATTTGGGCATCAATGCCGTGGCCTTACTACCGCGCTCTTGACGTGTCCGTCA 2928
QY 1202 TCCCGACCAAGTGTGACGTTGTCTGTGSCAACTGACGCCCTCATGACCGGCTTTACCG 1261
Db |||||
QY 2929 TCCCGACCAAGCGGAGATGTTGTCTGTGGCAACCGATGCCCTCATGACCGGCTATACCG 2988
QY 1262 GCACCTTCGATTCGGTGATAGACTGCAACAACGTTGTGTCAACCCAGACAGTCGACTTCAGCC 1321
Db |||||
QY 2989 GCACCTTCGACTCGGTGATAGACTGCAATAACGTTGTGTCAACCCAGACAGTCGACTTCAGCC 3048
QY 1322 TTGACCCCTACCTTCAACATTTGAGACAAATCAAGCTTCCCGAGGATGTGTCTCCCGTACTC 1381
Db |||||
QY 3049 TTGACCCCTACCTTCAACATTTGAGACAAATCAAGCTTCCCGAGGATGTGTCTCCCGCACTC 3108
QY 1382 AACGTGGGGTAGACTGGCAGAGGAAGCCAGGCATCTACAGATTTGTGGCAACCGGGG 1441
Db |||||
QY 3109 AACGTGGGGCAGACTGGCAGGGGAGACAGGCATCTACAGATTTGTGGCAACCGGGG 3168
QY 1442 AGCGTCTCTTCGCACTGTTTACTCGTCTGTCTCTCGAGTGTCTATGACGCGGTTGTG 1501
Db |||||
QY 3169 AGCGCCCTTCGGCATGTTTCGACTCGTCTCGTCTCTGTGAGTGTCTATGACGAGGCTGTG 3228
QY 1502 CTTTGTATGAGCTTACGCCCGCGAGACACAGTTAGGCTACGAGCATACATGAACACCC 1561
Db |||||
QY 3229 CTTTGTATGAGCTACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACCC 3288
QY 1562 CGGGACTTCCCGTGTGCAAGACCATCTTGAATTTTGGGAGGCGCTTTTACGGGTCTCA 1621
Db |||||
QY 3289 CGGGGCTTCCCGTGTGCGCAGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCTCA 3348
QY 1622 CCCACATAGACGCCCACTTCTCTATCCACAGCAAGCAGAGTGGGGAACCTTCCCTATC 1681
Db |||||
QY 3349 CTCATATAGATGCCCACTTCTCTATCCACAGCAAGCAGAGTGGGGAACCTTCTTACC 3408
QY 1682 TGTAGGCTACCAAGCCACCGTGTGCGCTAGACTCAAGCCCTCCCGCTCGTGGACC 1741
Db |||||
QY 3409 TGTAGGCTACCAAGCCACCGTGTGCGCTAGGCTCAAGCCCTCCCGCTCGTGGACC 3468
QY 1742 AGATGTGGAAGTGTCTTGCATCCGCTCAAGCCCAACCTTCCATGGGCCAACAACCTCTGCTAT 1801
Db |||||
QY 3469 AGATGTGGAAGTGTCTTGCATCCGCTCAAGCCCAACCTTCCATGGGCCAACAACCTCTGCTAT 3528
QY 1802 ATAGACTGGCGCTGTCCAGATGAAGTCAACCTGACGCAACCCAGTCAACCAAGTATATCA 1861
Db |||||
QY 3529 ACAGACTGGCGCTGTTTCAAGTGAATCAACCTGACGCAACCCAGTCAACCAAAATACATCA 3588
QY 1862 TGACATGTATGTGGCTGACTGGAGGTCTGTACGAGTACCTGGTGTCTGTTGGGGCG 1921
Db |||||
QY 3589 TGACATGTATGTGGCGCACTGTGAGGTCTGTACGAGCACTGGTGTCTGTTGGGGCG 3648
```

```
QY 1922 TTCTGCTGCTTTTGGCGCGTATTTCCTATCCACAGGTCGCTGTGTCATAGTAGGTAGGA 1981
Db |||||
QY 3649 TCCTGCTGCTTTTGGCGCGTATTTCCTATCCACAGGTCGCTGTGTCATAGTAGGTAGGA 3708
QY 1982 TTGCTCTTGTCCGGAAGCGGCAATCATACCCGACAGGAAGTCTCTTACCGGAGTTG 2041
Db |||||
QY 3709 TCGTCTTGTCCGGAAGCGGCAATCATACCTGACAGGAAGTCTCTTACCGAGAGTTG 3768
QY 2042 ATGAAATGGAAGAGTGCT 2059
Db |||||
QY 3769 ATGAGATGGAAGAGTGCT 3786
QY 3769 ATGAGATGGAAGAGTGCT 3786

RESULT 9
AAAT5296
ID AAAT5296 standard; cDNA; 8316 BP.
XX
AC AAAT5296;
XX
DT 15-JAN-2001 (first entry)
XX
DE cDNA sequence compiled Hepatitis C virus cDNA clones.
XX
KW Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;
XX
OS viral infectivity; viral replication; ds.
XX
FH Hepatitis C virus.
XX
Key Location/Qualifiers
FT 1.8316
FT /tag= a
FT /note= "partial sequence; no termination codon given"
XX
PN EP1034785-A2.
XX
PD 13-SEP-2000.
XX
PF 16-MAR-1990; 2000BP-00109602.
XX
PR 17-MAR-1989; 89US-00325338.
PR 20-APR-1989; 89US-00341334.
PR 18-MAY-1989; 89US-00355002.
PR 16-MAR-1990; 90EP-00302866.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Houghton M, Choo Q, Kuo G;
XX
DR WPI; 2000-566891/53.
XX
DR P-PSDB; AAB18540.
XX
PT Novel composition comprising a hepatitis C virus antisense polynucleotide
PT which is complementary to or corresponds to a sense strand of the virus
PT genome, and selectively hybridizes to it.
XX
PS Example; Fig 16; 75pp; English.
XX
CC The specification describes a pharmaceutical composition which comprises
CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
CC characterized by a positive stranded RNA genome which has 40% homology at
CC the polypeptide level to a HCV polyprotein. The antisense polynucleotide
CC binds to cellular polynucleotides which enhance and/or are required for
CC viral infectivity, replicative ability or chronicity. The antisense
CC polynucleotides may also be designed to bind with high specificity, to be
CC of increased stability, to be stable and to have low toxicity. The
CC composition also comprises an agent which causes viral RNA to be
CC inactive. The composition is used for preventing HCV replication in a
CC system. The present sequence represents a novel HCV cDNA sequence, which
CC is used in the course of the invention
XX
SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;
Query Match 86.7%; Score 1786; DB 3; Length 8316;
```


ID	AAZ07656 standard; DNA; 9133 BP.	QY	62	GCTTGCAGCGCGGACAAAAACAGGTGAGGGTTCAGATCGTGTCAACTGCTG	121
XX	AC	DB	3404	GCCTAACTGGCGGGACAAAAACAGGTGAGGGTTCAGATCGTGTCAACTGCTG	3463
XX	AAZ07656;				
DT	20-MAR-2003 (revised)	QY	122	CCAGACTTTCTTGGCAACTGATTAACGGGGTGTGTTGGACTGCTACCATGAGCCG	181
DT	08-NOV-1999 (first entry)	DB	3464	CCCAAACTTCTGGCAAGTGCATCAATGGGGTGTGCTGGACTGCTACCAAGGGCCG	3523
XX	Nucleotide sequence of HCV-1 ORF.				
XX	Hepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH;	QY	182	GAACAGGACCACTTGGTGCACCTTAAGGGTCTCTGTATCCAGATGTACACCAATGTGCAC	241
KW	HCV infection; vaccine; db.	DB	3524	GAACAGGACCACTGCGCTCACCAGGGTCTCTGTATCCAGATGTATACCAATGTGAGCC	3583
XX	Hepatitis C virus.	QY	242	AAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCGCGCTCATTAACACCATGCACTTGGC	301
XX	Key	DB	3584	AAGACCTTGTGGCTGGCCCGCTCCCAAGGTGAGCGCTCATTAACACCATGCACTTGGC	3643
FT	CDS				
FT	Location/Qualifiers				
FT	268..9132				
FT	/*tag= b				
FT	/transl_except= (pos:1588..1589; aa:Leu)				
FT	/note= "this codon has an apparent 1 nucleotide deletion,				
FT	which alters the reading frame"				
FT	/transl_except= (pos:1647..1650; aa:Pro)				
FT	/note= "this codon has an apparent 1 nucleotide				
FT	insertion, which alters the reading frame; this insertion				
FT	is not indicated in the sequence present in the formal				
FT	sequence listing of the specification"				
XX	EP939128-A2.	QY	422	GAGGCCCTCTGTGTGCCCCGAGGACATGCGGTAGGCATATTTCAGAGCGCGGTATGCA	481
XX		DB	3764	GGGTCCGCTGTGTGCCCCGAGGACATGCGGTAGGCATATTTCAGAGCGCGGTATGCA	3823
XX	01-SEP-1999.	QY	482	CCCGTGGAGTGGCTAAGGGCGTGGACTTCATCCCGTAGAGAGCTTAGAGACCAACATGA	541
XX	17-SEP-1990; 99EP-00101746.	DB	3824	CCCGTGGAGTGGCTAAGGGCGTGGACTTCATCCCGTAGAGAGCTTAGAGACCAACATGA	3883
PR	15-SEP-1989; 89US-00408045.	QY	542	GGTCCCGGTGTCTCAGACAACTCTCCGCCACACGAGCTGCCACAGACTACCAAGTGG	601
PR	21-DEC-1989; 89US-00456142.	DB	3884	GGTCCCGGTGTCTCAGACAACTCTCTCCACAGTAGTGCCCCAGAGCTTCCAGGTGG	3943
PR	17-SEP-1990; 90EP-00310149.				
XX	(OYAA/) OYA A.	QY	602	CCCACCTGCATGCTCCACCGGCGGTAGAGACCAAGGTGCCCGCGCGCATAGCGAG	661
PA	(CHIR) CHIRON CORP.	DB	3944	CTCACCTCCATGCTCCACAGGCGGCAAGACCAAGGTGCCCGCGCATAGCGAG	4003
XX	Miyamura T, Saito I, Houghton M, Weiner AJ, Han J, Kolberg JA;				
PI	Cha T, Irvine BD;	QY	662	CTCAGGGCTACAAGGTGTGTGCTCAACCCCTCGTTCGCTGCGCAACAATGGGGCTTGGTG	721
XX	WP1; 1999-480843/41.	DB	4004	CTCAGGGCTATAAGGTGTGTGCTCAACCCCTCTGTCTGCTGCAACACTGGGGCTTGGTG	4063
DR	P-PSDB; AAY14975.				
XX	New Hepatitis C Virus isolates, useful for diagnosis of hepatitis	QY	722	CTTACATGTCACAGGCCCATGGGATGTATCTTAACATCAGGACTGGGGTGAGGACAATTA	781
PT	infections and development of vaccines.	DB	4064	CTTACATGTCACAGGCCCATGGGATGTATCTTAACATCAGGACTGGGGTGAGGACAATTA	4123
XX	Disclosure; Fig 12; 132pp; English.	QY	782	CTACTGGCAGCGCGATCAGGTATTCCACCTACGCGCAAGTTCCTTGCCGACGGCGGTGT	841
PS		DB	4124	CCACTGGCAGCGCGATCAGGTATTCCACCTACGCGCAAGTTCCTTGCCGACGGCGGTGT	4183
CC	The invention provides two new isolates of hepatitis C virus (HCV), J1	QY	842	CAGGGGTGCTTATGACATAATAATTTGTGACAGGTGCCACTCCACGGATGCAACATCA	901
CC	and J7. These two isolates comprise nucleotide and amino acid sequences	DB	4184	CAGGGGTGCTTATGACATAATAATTTGTGACAGGTGCCACTCCACGGATGCAACATCA	4243
CC	that are distinct from the HCV isolate HCV-1. The nucleotide sequences	QY	902	TCCTGGGCAATTCGCACTGCTTGTGACAGGACCGCGGGCGGAGACTGACTGTGTC	961
CC	may be used to detect non-A, non-B HCV (NANBH) polynucleotides by	DB	4244	TCCTGGGCAATTCGCACTGCTTGTGACAGGACCGCGGGCGGAGACTGACTGTGTC	4303
CC	hybridization for diagnosis of NANBH infections. They may also be used to	QY	962	TGCGCACCCCTACCCCTCCGGGCTCCGTCATCTGTGCCCATCTTAACATCGAGGAGTTG	1021
CC	screen blood donors, donated blood and blood products for this infection.	DB	4304	TGCGCACCCCTACCCCTCCGGGCTCCGTCATCTGTGCCCATCTTAACATCGAGGAGTTG	4363
CC	The isolates may also be used to isolate other naturally occurring	QY	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGAAGGCTATTCCTTGAAGCAATTA	1081
CC	variants of the virus. The polypeptides may be used as a vaccine for	DB	4364	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGAAGGCTATTCCTTGAAGCAATTA	4423
CC	administration to patients to protect against infection with NANBH. The	QY	1082	AGGGGGGAGACATCTCATCTTCTCCACTAAAGAGAGTGGCAGGACTCGCGCA	1141
CC	present sequence represents the nucleotide sequence of HCV-1 ORF.	DB	4424	AGGGGGGAGACATCTCATCTTCTCCACTAAAGAGAGTGGCAGGACTCGCGCA	4483
CC	(Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to	QY	1142	AACTGTGCGGTGGGCGTCAATGCGGTGGCTTACTACCGCGGCTTGTGTGTCGTC	1201
CC	correct PR field.)				
XX	Sequence 9133 BP; 1834 A; 2772 C; 2600 G; 1927 T; 0 U; 0 Other;				
SQ					
Query Match	86.7%; Score 1786; DB 2; Length 9133;				
Best Local Similarity	91.7%; Pred. No. 0;				
Matches 1888; Conservative	0; Mismatches 170; Indels 0; Gaps 0;				
QY	2 TGCGCGCTATCAGCGCTATGCCAGCAGCAGAGGGCGCTTTGGGATGCATATACCA				61
DB	3344 TGCGCGCGCTATCAGCGCTATGCCAGCAGCAGAGGGCGCTTTGGGATGCATATACCA				3403

```
Db 4484 AGCTGGTCGCAATGGGCAATCAATCGCGTGGCTACTACCGGGTCTTGAAGTGTCCGTC 4543
Qy 1202 TCCGACACAGTGTGACGTTGCTCGTGGCAACTGACCGCTCATGACCGGCTTTACCG 1261
Db 4544 TCCGACACAGGCGATGTTGCTCGTGGCAACCGATGCCCTCATGACCGGCTATACCG 4603
Qy 1262 GCGACTTCGATTCGGTGTGATGACCTGCAACACGTTGTGACCCAGACAGTGTGACGCC 1321
Db 4604 GCGACTTCGATTCGGTGTGATGACCTGCAATCGTGTGACCCAGACAGTGTGATTCAGCC 4663
Qy 1322 TTGACCTTACCTTACCATGAGACATCAGCTTCCCGAGGATGCTCTCCCGTACTC 1381
Db 4664 TTGACCTTACCTTACCATGAGACATCAGCTTCCCGAGGATGCTCTCCCGCACTC 4723
Qy 1382 AACCTCGGGTAGGACTGGCAGAGGAAGCAGGACATCTACAGATTTGTGGCACCAGGGG 1441
Db 4724 AACCTCGGGTAGGACTGGCAGAGGAAGCAGGACATCTACAGATTTGTGGCACCAGGGG 4783
Qy 1442 AGCTGCTTTCGGCATGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
Db 4784 AGCGCCCTCCCGCATGTTGACTCGTCCGTCCTGTGAGTGTGATGAGCGAGGCTGTG 4843
Qy 1502 CTTGGTATGACTTACCGCCCGCAGACACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1561
Db 4844 CTTGGTATGACTTACCGCCCGCAGACACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4903
Qy 1562 CGGACTTCCCGTGTGCAAGACCATCTTGAATTTTGGAGGGCTCTTTACGGGTCTCA 1621
Db 4904 CGGGGCTTCCCGTGTGCAAGACCATCTTGAATTTTGGAGGGCTCTTTACAGGCTCA 4963
Qy 1622 CCCATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
Db 4964 CTCATATAGTCCCACTTCTTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC 5023
Qy 1682 TGCTAGGCTACCAAGCACCGTGTGCGTAGAGCTCAAGCCCTCCCGCTGCTGGGACC 1741
Db 5024 TGCTAGGCTACCAAGCACCGTGTGCGTAGAGCTCAAGCCCTCCCGCTGCTGGGACC 5083
Qy 1742 AGATGTGGAAGTGTGATCGCTCTCAAGCCACCTTCCATGGGCCCAACACCTCTGCTAT 1801
Db 5084 AGATGTGGAAGTGTGATCGCTCTCAAGCCACCTTCCATGGGCCCAACACCTCTGCTAT 5143
Qy 1802 ATAGACTGGGCGTGTCCAGATGAAGTCAACCTTGACGACCCAGTCAACCAAGTATATCA 1861
Db 5144 ACAGACTGGGCGTGTTCAGATGAAGTCAACCTTGACGACCCAGTCAACCAATACATCA 5203
Qy 1862 TGACATGTATGTGGCTGTGACCTGAGTGTGACAGTACCTGGGTGCTGTGGCGGCG 1921
Db 5204 TGACATGTATGTGGCGGACCTGGAGGTGCTGACGAGCACCTGGGTGCTGTGGCGGCG 5263
Qy 1922 TTCTGGCTGTCTTGGCGCGTATGCTTATCCAGAGCTGCTGCTCATAGTAGGTAGGA 1981
Db 5264 TCCTGGCTGTCTTGGCGCGTATGCTTATCCAGAGCTGCTGCTCATAGTAGGTAGGA 5323
Qy 1982 TTGTCTTGTCCGAAAGCCGCAATCATACCCGACGGGAAGTCTCTACCGGAGTTCG 2041
Db 5324 TCGTCTTGTCCGAAAGCCGCAATCATACCTGACAGGAAGTCTCTACCGAGTTCG 5383
Qy 2042 ATGAATGGAAGTGTCT 2059
Db 5384 ATGAGATGGAAGTGTCT 5401
```

RESULT 11

AAQ10566 standard; DNA; 9185 BP.

XX

AC AAQ10566;

XX

DT 25-MAR-2003 (revised)

DT 29-APR-1991 (first entry)

XX

DE Hepatitis C virus strain 1 DNA.
XX
KW Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;
KW viral infections; ss.
XX
OS Hepatitis C virus.
XX
XX EP414475-A.
XX
XX 27-FEB-1991.
XX
XX 21-AUG-1990; 90EP-00309120.
XX
XX 25-AUG-1989; 89US-00398667.
XX
XX (CHIR) CHIRON CORP.
XX
XX Weiner AJ, Steimer KS;
PI
XX WPI; 1991-059670/09.
XX
XX Cell lines infected with hepatitis C virus - are used as source of
PT antigens for detection of HCV antibodies, for vaccines, and for screening
PT anti-viral agents.
XX
XX PS Disclosure; Fig 1; 24pp; English.
XX
CC This is a hepatitis C virus (HCV) composite cDNA sequence, deduced using
CC overlapping clones. A compsn. contg. the antigenic protein encoded by
CC this sequence is useful for detecting anti-HCV anti- bodies (Abs) and for
CC screening an agent which inhibits HCV replic- ation. A cell line infected
CC with this virus can be used as a source of antigens. The antigen is
CC useful for preparing vaccines for treating viral infections. See also
CC AAQ10567. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Query Match 86.7%; Score 1786; DB 2; Length 9185;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

```
Qy 2 TGGCGCTATCAGCGCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCAATATCACCA 61
Db 3396 TGGCGCGCATACGCGGTACGCCAGCAGACAAAGGGGCTCTTAGGGTGCATATCACCA 3455
Qy 62 GCTTGACCGCGCGGACAAACCCAGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG 121
Db 3456 GCCTAATGCGCGGACAAACCCAGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG 3515
Qy 122 CCAGAGCTTTCTTTGGCAACCTGCATTAACCGGGGTGTGTTGGACTGTCTACCATGGAGCCG 181
Db 3516 CCANACCTTCTTGGCAACGTGCATCAATGGGGTGTGCTGACTGTCTACCGGGGCGG 3575
Qy 182 GAACAAGGACCATTTGGTCACTTAAGGGTCTCTTATFCCAGATGTACACCAATGTGACC 241
Db 3576 GAACGAGGACCATCGGTTCACCAAGGGTCTCTGCATCCAGATGTATACCAATGTAGACC 3635
Qy 242 AAGACCTCGTAGGCTGGCGGCTCCCAAGGTGCGGCTCATTAACACCATGCACTTGGCG 301
Db 3636 AAGACCTTGTGGGCTGGCGGCTCCCGAGGTAGCGGCTCATTAACACCATGCACTTGGCG 3695
Qy 302 GCTCTCTGGACCTTTTACCTGTCACGAGGCACCGCATGTCAATTCCTGTGCGCGACGGG 361
Db 3696 GCTCTCTGGACCTTTTACCTGTCACGAGGCACCGCATGTCAATTCCTGTGCGCGGCGG 3755
Qy 362 GTGATGGCAGGGCAGCGCTGCTTTTCGCCCGGCTATTCCTTACTTGAAGGCTCTCGG 421
Db 3756 GTGATAGCAGGGCAGCGCTGCTGCGCCCGGCCCATTTCTACTTTGAAGGCTCTCTCGG 3815
Qy 422 GAGGCGCTGCTGCTGCGCCCGCAGGACATGCCCTAGGCATATTCAGAGCGCGGTATGCA 481
Db 3816 GGGGTCGCTGTTGTGCGCCCGCGGCGACCGCTGGGCAATATTTAGGGCCCGCGGTATGCA 3875
```

Qy	482	CCCGTGGAGTGGCTAAAGCGGTGGAATTCATCCCGGTAGAGAGCTTAGAGACAACCAATGA	541
Db	3876	CCCGTGGAGTGGCTAAAGCGGTGGAATTCATCCCGTGGAGAACTTAGAGACAACCAATGA	3935
Qy	542	GGTCCC CGGTGTTCTCAGACAATCTCCCCACACAGCAGTGCSCCCAGAGCTACCAAGTGG	601
Db	3936	GGTCCC CGGTGTTTACGGATAACTTCTCTCCACAGTAGTGCCCCAGAGCTTCCAGGTGG	3995
Qy	602	CCCACTGTCATGCTCCCAACCGCAGCGGTAAAGACACCAAGTCCCGCGCGCATACGCAG	661
Db	3996	CTCACTCCATGCTCCCAACAGCAGCGGCAAAAGACCAAGTCCCGCTGCATATGCAG	4055
Qy	662	CTCAGGGCTACAAGGTGCTGTGTCTCAACCCCTCGTGTGTGTGCAACAATGGCGCTTGGTG	721
Db	4056	CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTTGGTG	4115
Qy	722	CTTACATGTCCAAGGCCATGGGATTCATCTTAACATCAGGACTGGGGTGAGGACAATTA	781
Db	4116	CTTACATGTCCAAGGGCTATGGGATCGATCTCTAAACATCAGGACCGGGTGAGAACAAATTA	4175
Qy	782	CTACTGGCAGCCCGATCACGTATATCCACCTACGGCAAGTTCCTTTGCCGACGGCGGGTGT	841
Db	4176	CCACTGGCAGCCCATCACGTACTCCACCTACGGCAAGTTCCTTTGCCGACGGCGGGTGT	4235
Qy	842	CAGGGGTGTTATGACATAATAATTTGTGACAGTGCCTCACCGATGCAACATPCCA	901
Db	4236	CGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCTCACCGATGCAACATPCCA	4295
Qy	902	TCITGGGCNTGGCACTGTCCTTGACCAAGCAGAGACCGGGGGCGAGACTGACTGTGC	961
Db	4296	TCITGGGCATCGGCACCTGTCCTTGACCAAGCAGAGACTGGGGGGCGAGACTGCTGTGTGC	4355
Qy	962	TCGGCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCCACTCTTAACATCGAGGAGTTG	1021
Db	4356	TCGGCACCGGCACCCCTCCGGGCTCCGTCACTGTGCCCCACTCCCAACATCGAGGAGTTG	4415
Qy	1022	CTCTGTCCACTPACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCTTTGAAGCAATTA	1081
Db	4416	CTCTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATCA	4475
Qy	1082	AGGGGGGAGACATCTCATCTTGTGCCACTCAAGAGAGAGTGGCAGCAGCTCGCCGCAA	1141
Db	4476	AGGGGGGAGACATCTCATCTTGTTCATTCAAAGAGAGTGGCAGCACTCGCCGCAA	4535
Qy	1142	AACTGTTCCGCTTGGGCGTCAATGCCGTGGCTTACTACCGGGCTTGAATGTTCGTCTCA	1201
Db	4536	AGCTGTTCCGATTTGGGCATCAATGCCGTGGCTTACTACCGGGCTTGAATGTTCGTCTCA	4595
Qy	1202	TCCCGACCAAGTGTGACGTTGTGTGTGTGCAACTGACGCCCTCATGACCGGGCTTACCG	1261
Db	4596	TCCCGACCAAGCGGATGTTGTGTGTGTGCAACCGATGCCCTCATGACCGGCTATACCG	4655
Qy	1262	CGGACTTCGATTCCGGTGATGACTGCAACACGTTGTGTACCCGACAGTGCATTCAGCC	1321
Db	4656	CGCACTTCGATCCGGTGATGACTGCAATACGTTGTGTACCCGACAGTGCATTCAGCC	4715
Qy	1322	TTGACCTTACTCTCACCATTTGAGACAATCAGCTTCCCAAGGATGCTGTCTCCGTAATC	1381
Db	4716	TTGACCTTACTCTTACCATTGAGACAATCAGCTTCCCAAGGATGCTGTCTCCGCAATC	4775
Qy	1382	AACTCGGGGTAGGACTGGCAGAGGGAAGCCAGCATCTACAGATTTTGTGGCACCCGGGG	1441
Db	4776	AACTCGGGGCAGGACTGGCAGGGGAAGCCAGCATCTACAGATTTTGTGGCACCCGGGG	4835
Qy	1442	AGGTCCTTCTGGGATGTTTGAATCTGTTCTCTCTCGGAGTGCTATGACCGCGGTTGTG	1501
Db	4836	AGCGCCCTTCGGCATGTTTCGACTCTGTCGTCTCTCTGAGTGTATATGACGAGGCTGTG	4895
Qy	1502	CTTGTGATGACTTTACCGCCCGCAGACACAGTTAGGCTTACGAGCATATCAACAACCC	1561
Db	4896	CTTGTGATGACTTACCGCCCGCAGACTACAGTTAGGCTTACGAGCGGTAATGAACAACCC	4955
Qy	1562	CGGACTTCCCGTGTGCAAGAACCATCTTGAAATTTTTGGAGGGCGCTTTTACCGGGTCTCA	1621

Db	4956	CGGGCGCTTCCGCTGTGCAGGACCATTTGAATTTTGGGAGGCGCTTTTACAGGCGCTCA	5015
Qy	1622	CCACATAGACGCCCACTTCTATCCAGACAAGCAGAGTGGGGAACAACTTCCCTATC	1681
Db	5016	CTCATATAGATGCCCACTTTCTATCCAGACAAGCAGAGTGGGAGAACCTTCTTACC	5075
Qy	1682	TGCTAGCGTACCAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC	1741
Db	5076	TGGTAGCGTACCAGCCACCGTGTGCGCTAGGAGCTCAAGCCCTCCCGCATCGTGGGACC	5135
Qy	1742	AGATGTGGAAGTGGTGTGATCCGCTCTAAGCCCACTTCATAGGGCCAAACCTCTGCTAT	1801
Db	5136	AGATGTGGAAGTGGTGTGATTCGCTCTAAGCCCACTTCATAGGGCCAAACCTCTGCTAT	5195
Qy	1802	ATAGACTGGCGCGCTGTCAGAAATGAAGTCAACCCTGACGACCCAGTCACCAAGTATATCA	1861
Db	5196	ACAGACTGGCGCGCTGTTCAGAAATGAAGTCAACCCTGACGACCCAGTCACCAATACATCA	5255
Qy	1862	TGACATGATATGTCCGGCTGACCTGGAGGTGCTCACGAGTACCTGGGTGCTGTTGGCGGCG	1921
Db	5256	TGACATGATATGTCCGGCGGCTGAGGTGCTCACGAGCACCTGGGTGCTGTTGGCGGCG	5315
Qy	1922	TTCTGGCTGCTTTGGCGCGCTATGCTTATCCACAGGTGCGGTGCTCATAGTAGGTAGGA	1981
Db	5316	TCCTGGCTGCTTTGGCGCGCTATGCTTATCCACAGGTGCGGTGCTCATAGTAGGTAGGA	5375
Qy	1982	TTGTCTTGTCCGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTGCG	2041
Db	5376	TCGTCTTGTCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGGAGTTGCG	5435
Qy	2042	ATGAAATGGAAGAGTGCT	2059
Db	5436	ATGAGATGGAAGAGTGCT	5453
RESULT 12			
ID	AAA75297		
DD	AAA75297 standard; cdna; 9185 BP.		
XX	AC AAA75297;		
XX	15-JAN-2001 (first entry)		
XX	Sense strand of HCV encoding a polyprotein.		
XX	Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;		
XX	viral infectivity; viral replication; ds.		
XX	Hepatitis C virus.		
XX	Key	Location/Qualifiers	
PH	CDS	320..9184	
FT	/*tag= a		
FT	/note= "partial sequence; no termination codon given"		
XX	EPI034785-A2.		
XX	13-SEP-2000.		
XX	16-MAR-1990; 2000EP-00109602.		
XX	17-MAR-1989; 89US-00325338.		
XX	20-APR-1989; 89US-00341334.		
XX	18-MAY-1989; 89US-00355002.		
XX	16-MAR-1990; 90EP-00302866.		
XX	(CHIR) CHIRON CORP.		
XX	Houghton M, Choo Q, Kuo G;		
XX	WPI; 2000-566891/53.		
XX	P-PSDB; AAB18541.		

XX Novel composition comprising a hepatitis C virus antisense polynucleotide
PT which is complementary to or corresponds to a sense strand of the virus
PT genome, and selectively hybridizes to it.

XX Example; Fig 17; 75pp; English.

CC The specification describes a pharmaceutical composition which comprises
CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
CC characterized by a positive stranded RNA genome which has 40% homology at
CC the polypeptide level to a HCV polypeptide. The antisense polynucleotide
CC binds to cellular polynucleotides which enhance and/or are required for
CC viral infectivity, replicative ability or chronicity. The antisense
CC polynucleotides may also be designed to bind with high specificity, to be
CC of increased stability, to be stable and to have low toxicity. The
CC composition also comprises an agent which causes viral RNA to be
CC inactive. The composition is used for preventing HCV replication in a
CC system. The present sequence represents a novel HCV cDNA sequence, which
CC is used in the course of the invention

XX Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Query Match 86.7%; Score 1786; DB 3; Length 9185;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy	2	TGGCGCCTATCACGGCCTATGCCAGCAGACAAGGGGCTTTTGGGATGCATATACCA	61
Db	3396	TGGCGCCATCACGGCGTACGCCAGCAGACAAGGGGCTTCTAGGGTGCATATACCA	3455
Qy	62	GCTTAGCGCGGGGCAAAAACACAGATGGAGGTTCAGATGTCATCTGCTG	121
Db	3456	GCCTAACTGGCGGGGCAAAAACCAAGTGGAGGTTCAGATGTCATCTGCTG	3515
Qy	122	CCAGAGCTTCTTGGCACTGCAATTAACGGGTGTTGGACTGTTACCATGGAGCG	181
Db	3516	CCCAAACTTCTTGGCACTGCAATTAATGGGTGTTGGACTGTTACCATGGAGCGG	3575
Qy	182	GAACAAGGACCATTTGCGTCACTAAGGCTCCTGTTATCCAGATGTACACCAATGTGACC	241
Db	3576	GAACAGGACCATTCGCTCACTAAGGCTCCTGTTATCCAGATGTATACCAATGTGACC	3635
Qy	242	AAGACCTGTAGGTGGCGGCTCCCAAGGTGCCGCTCATTAACCAATGTGCTTGG	301
Db	3636	AAGACCTGTGGGTGGCGGCTCCGCAAGGTAGCGCTCATTAACCAATGTGCTTGG	3695
Qy	302	GCTCCTCGGACCTTACCTGTGACAGGACAGCGGATGTCATTCCTGTGGCGGACGG	361
Db	3696	GCTCCTCGGACCTTACCTGTGTCAGAGGACAGCGGATGTCATTCCTGTGGCGGCGG	3755
Qy	362	GTGATGGCAGGGCAGCTGCTTTTCGCCCGGCTTCTCTTACTTGAAAGGCTCCTCGG	421
Db	3756	GTGATAGAGGGGACGCTGCTGTGCGCCGCGCCATTTCTTCTTGAAGGCTCCTCGG	3815
Qy	422	GAGGCCCTCTGCTGTGCCCGCAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGCA	481
Db	3816	GGGGTCCGCTGTGTGCCCCCGGGGACGCGGTGGGCAATTTTAGGGCCGCGGTGCA	3875
Qy	482	CCCGTGAAGTGGCTTAAGCGGTGACCTATTCCTCCGCTAGAGAGCTTAGAGAACCATGA	541
Db	3876	CCCGTGAAGTGGCTTAAGCGGTGACCTATTCCTCTGTGGAAACCTTAGAGAACCATGA	3935
Qy	542	GGTCCCGGCTTCTCAGACAACTCTCCACAGCAGTGCCTCCAGAGTACCAAGTGG	601
Db	3936	GGTCCCGGCTTCTCAGGATACTCTCTCCACAGTAGTGCCTCCAGAGTTCAGAGTGG	3995
Qy	602	CCCACTGCTATGCTCCCAACCGGACGGGTAGAGCAACAAGGTCCCGCGCATACGACG	661
Db	3996	CTCACTCCATGCTCCCAACCGGACGGGTAGAGCAACAAGGTCCCGCGCTGATATGAG	4055
Qy	662	CTCAGGCTCAAGGCTGCTGCTCAACCGCTCCGTTGCTGCAACAATGGGCTTTGGTG	721
Db	4056	CTCAGGCTCAAGGCTGCTGCTCAACCGCTCTGTTGCTGCAACAATGGGCTTTGGTG	4115

Qy	722	CTTACATGTCCAAAGCCCATGGGATTCATCTAAACATCAGGACTGGGGTGAGGCAATTA	781
Db	4116	CTTACATGTCCAAAGCTCATGGGATTCATCTAAACATCAGGACCGGGGTGAGAACATTA	4175
Qy	782	CTACTGGCAGCCGATCACGTTATTCACCTACGGAAGTTCTTTCCGACCGGGGTGTT	841
Db	4176	CCACTGGCAGCCCATCACGTTATTCACCTACGGAAGTTCTTTCCGACCGGGGTGTT	4235
Qy	842	CAGGGGTGCTTATGACATATAATTTGTACGAGTGCCTCCACGGATGCAACATCCA	901
Db	4236	CGGGGGGCTTATGACATATAATTTGTACGAGTGCCTCCACGGATGCAACATCCA	4295
Qy	902	TCTTGGCATTTGGCACTGCTCTTGAACCAAGCAGAGACCGGGGGCGAGACTGATGTC	961
Db	4296	TCTTGGCATTCGGCACTGCTCTTGAACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTC	4355
Qy	962	TCGCCACCGTACCCCTCCGGGCTCCGTCACCTGTGCGCCATCTTAACATCAGAGAGTTG	1021
Db	4356	TCGCCACCGCACCCCTCCGGGCTCCGTCACCTGTGCGCCATCTTAACATCAGAGAGTTG	4415
Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCTTCTTGAACCAATTA	1081
Db	4416	CTCTGTCCACACCGGAGAGATCCCTTTTATGCAAGGCTATTCCTTCTTGAAGTATCA	4475
Qy	1082	AGGGGGGAGACATCTCTCTTCTGCACTCAAAAGAAAGTGCAGAGCTCGCCGCA	1141
Db	4476	AGGGGGGAGACATCTCTCTCTGTCATTAAGAAAGTGCAGAGCTCGCCGCA	4535
Qy	1142	AACGTGCGGTGGGGTCAATGCGGTGTTACTACCGGGCTTTGATGTGCTGCTCA	1201
Db	4536	AGCTGTGCGATTTGGGATCAATGCGGTGCTACTACCGGGTCTTGAAGCTGCTGCTCA	4595
Qy	1202	TCCGACACGAGTGTGCTGCTGCGCACTGACGCGCTCATCAGCGGCTTTACCG	1261
Db	4596	TCCGACACGAGGATGTTGCTGCTGCGCACTGACGCGCTCATCAGCGGCTTTACCG	4655
Qy	1262	GCAGCTTCGATTCGGTGATAGACTGCAACACGCTGTGTCACCCAGACAGTGCATTCAGCC	1321
Db	4656	GCAGCTTCGATTCGGTGATAGACTGCAATAGCTGTGTCACCCAGACAGTGCATTCAGCC	4715
Qy	1322	TTGACCTTACCTTACCATTTGAGACATACGCTTCCCGAGGATGCTGTCTCCGTA	1381
Db	4716	TTGACCTTACCTTACCATTTGAGACATACGCTTCCCGAGGATGCTGTCTCCCGACTC	4775
Qy	1382	AACGTGCGGTAGGACTGGCAGAGGAGAGCAGGATCTACAGATTTGTGCGACCGGGG	1441
Db	4776	AACGTGCGGTAGGACTGGCAGAGGAGAGCAGGATCTACAGATTTGTGCGACCGGGG	4835
Qy	1442	AGCGTCTCTTGGCATGTTTGAATCTGCTCTGCGAGTGTATGACGCGGTTGTTG	1501
Db	4836	AGCGCTCTCTGGCATGTTTGAATCTGCTCTGCGAGTGTATGACGCGGTTGTTG	4895
Qy	1502	CTTTGATGAGCTTACCGCCCGCAGACCAAGTTAGGCTACGAGCATATCATGAACACC	1561
Db	4896	CTTTGATGAGCTTACCGCCCGCAGACCAAGTTAGGCTACGAGCGTACATGAACACC	4955
Qy	1562	CGGCACTTCCCGTGTGCAAGACATCTTGAATTTTGGGAGGCGCTTTTACGGGCTCA	1621
Db	4956	CGGCGCTTCCCGTGTGCAAGACATCTTGAATTTTGGGAGGCGCTTTTACAGGCTCA	5015
Qy	1622	CCACATAGACGCCCTTCTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC	1681
Db	5016	CTCATATGATGCCCTTCTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC	5075
Qy	1682	TGCTAGGCTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTTCCCGCTCGTGGGACC	1741
Db	5076	TGCTAGGCTACCAAGCCACCGTGTGGCTAGGCTCAAGCCCTTCCCGCTCGTGGGACC	5135
Qy	1742	AGATGGGAAGTGTGATCCGCTCAAGCCCACTTCCATGGGCCAACCTTCTGTAT	1801
Db	5136	AGATGGGAAGTGTGATCCGCTCAAGCCCACTTCCATGGGCCAACCTTCTGTAT	5195

QY 1802 ATAGACTGGCGGTGTCAGAAATGAAGTCAACCTGACGACACCCAGTCAACCAAGTATATCA 1861
DB |||||
5196 ACAGACTGGCGGTGTCAGAAATGAAGTCAACCTGACGACACCCAGTCAACCAATATATCA 5255
QY 1862 TGACATGTATGTGCGGTGACCTGGAGTGTCTACAGATACCTGGGTGCTGCTTTGGCGGG 1921
DB |||||
5256 TGACATGTATGTGCGGTGACCTGGAGTGTCTACAGATACCTGGGTGCTGCTTTGGCGGG 5315
QY 1922 TTCTGGCTGCTTTGGCGGGTATTTGCTATCCACAGCTGCGTGTGTCATAGTAGGTAAGGA 1981
DB |||||
5316 TCTGGCTGCTTTGGCGGGTATTTGCTGTCAACAGCTGCGTGTGTCATAGTAGGTAAGGA 5375
QY 1982 TTGTCTTTGTCGGGAAGCCGGCAATCATATCCGACAGGAAGTCTCTACCGGAGTTGG 2041
DB |||||
5376 TCGTCTTTGTCGGGAAGCCGGCAATCATATCCGACAGGAAGTCTCTACCGAGAGTTGG 5435
QY 2042 ATGAATGGAAGTGTCT 2059
DB |||||
5436 ATGAGATGGAAGTGTCT 5453

RESULT 13

ADN35979

ID ADN35979 standard; cDNA; 9185 BP.

XX AC ADN35979;

XX DT 17-JUN-2004 (first entry)

XX HCV cDNA clone #2.

XX KW Antiviral; Vaccine; hepatitis C virus infection; HCV infection; ss.

XX OS Hepatitis C virus.

XX PN EP1394255-A2.

XX PD 03-MAR-2004.

XX PF 16-MAR-1990; 2003EP-00016585.

XX PR 17-MAR-1989; 89US-00325338.

XX PR 20-APR-1989; 89US-00341334.

XX PR 18-MAY-1989; 89US-00355002.

XX PR 16-MAR-1990; 90EP-00302866.

XX PA (CHIR) CHIRON CORP.

XX PI Houghton M, Choo Q, Kuo G;

XX DR WPI; 2004-193149/19.

XX DR F-PSDB; ADN35978.

XX Novel purified hepatitis C virus polypeptide comprising epitope encoded
PT by hepatitis C virus cDNA, useful as vaccine for treating hepatitis C
PT virus.

XX PS Claim 1; Fig 17; 79pp; English.

XX The present invention relates to hepatitis C virus (HCV) proteins and
CC cDNA sequences. The sequences are useful in immunoassays for detecting
CC antibodies directed against HCV antigen; preparing host cells transformed
CC with a recombinant polynucleotide; screening antiviral agents and
CC determining the effect of antiviral agent in inhibiting viral replication
CC in cell culture system; and developing vaccine for treating HCV
XX infection.

XX SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Query Match 86.7%; Score 1786; DB 12; Length 9185;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTATACGGCCTATGCCAGCAGCAAGGGGCCCTTTTGGGATGCATTAATCACCA 61
DB |||||
3396 TGGCGCCTATACGGCGGTATCGCCAGCAGCAAGGGGCCCTCTAGGGTGCATAATCACCA 3455
QY 62 GCTTGAACGGCGGGGCAAAAACCAAGGTGAGGGTGCAGTTCAGATCGTGTCAACTGCTG 121
DB |||||
3456 GCTTAATCGCGGGGCAAAAACCAAGTGGAGGGTGCAGTTCAGATCGTGTCAACTGCTG 3515
QY 122 CCAGACTTTTCTTGGCAACTGCTATTAACGGGTGTGTTGGAGTGTCTACCATGGAGCCG 181
DB |||||
3516 CCAGAACTTTCTGGCAACTGCTATTAACGGGTGTGTTGGAGTGTCTACCATGGAGCCG 3575
QY 182 GAACAGGACCAATTCGCTCACTTAAGGGTCTCTTTATCCAGATGTACACCAATGTGACC 241
DB |||||
3576 GAACAGGACCAATTCGCTCACTTAAGGGTCTCTTTATCCAGATGTATCAATGTAGACC 3635
QY 242 AAGACCTGTAGCTGGCGCGCTCCCAAGGTGCGCGCTCATTAACACCATGCACTTGG 301
DB |||||
3636 AAGACCTGTAGCTGGCGCGCTCCCAAGGTGCGCGCTCATTAACACCATGCACTTGG 3695
QY 302 GCTCTCTCGGACCTTTTACCTGGTCAAGGACGCGCGATGTCTCTGTGCGCGGACGGG 361
DB |||||
3696 GCTCTCTCGGACCTTTTACCTGGTCAAGGACGCGCGATGTCTCTGTGCGCGGCGGG 3755
QY 362 GTGATGCGAGGGGCGAGCTGTCTTCCCGCGCGCTATCTCTTACTTGAAGGCTCTCTCG 421
DB |||||
3756 GTGATGCGAGGGGCGAGCTGTCTTCCCGCGCGCTATCTCTTACTTGAAGGCTCTCTCG 3815
QY 422 GAGGCGCTCTGTGTCGCCCGCAGGACATGCGGTAGGCAATATTCAGAGCGCGGATGCA 481
DB |||||
3816 GAGGCGCTCTGTGTCGCCCGCAGGACATGCGGTAGGCAATATTCAGAGCGCGGATGCA 3875
QY 482 CCGCTGGAGTGGCTAAGCGGTGGACTTCATCCCGGTAGAGACTTAGAGCAACCAATGA 541
DB |||||
3876 CCGCTGGAGTGGCTAAGCGGTGGACTTCATCCCGGTAGAGACTTAGAGCAACCAATGA 3935
QY 542 GGTCCCGCGGTCTCTCAGACAACTCTCCCGCAGCAGTGCCTCCAGAGCTACCAAGTGG 601
DB |||||
3936 GGTCCCGCGGTCTCTCAGACAACTCTCTCCAGTAGTGCCTCCAGAGCTTCCAGGTGG 3995
QY 602 CCGACCTGATGCTCCACCGGAGGGTGAAGACCAAGTCCCGCGCGGATAGCGAG 661
DB |||||
3996 CTCACCTCCATGCTCCACAGGCGCGCAAGACCAAGGTCCCGCGGTCATATGCGAG 4055
QY 662 CTCAGGCTTACAAGGTGCTGCTCAACCCCTCTGCTGCAACAATGGGCTTTGGTG 721
DB |||||
4056 CTCAGGCTTACAAGGTGCTGCTCAACCCCTCTGCTGCAACAATGGGCTTTGGTG 4115
QY 722 CTTACATGTCCAAGGCCCATGGATGTATCTTAACATCAGGACTGGGGTGAGGACAATTA 781
DB |||||
4116 CTTACATGTCCAAGGCCCATGGATGTATCTTAACATCAGGACTGGGGTGAGGACAATTA 4175
QY 782 CTACTGGAGCCCGATCAGGTATTCACCTACGGCAAGTTCCTTGGCGAGCGGGGTGT 841
DB |||||
4176 CCACCTGGAGCCCGATCAGGTATTCACCTACGGCAAGTTCCTTGGCGAGCGGGGTGT 4235
QY 842 CAGGGGTGCTTATGACATAATAATTTGTGACAGTGCCACTCCAGGATGCAACATCCA 901
DB |||||
4236 CAGGGGTGCTTATGACATAATAATTTGTGACAGTGCCACTCCAGGATGCAACATCCA 4295
QY 902 TCTTGGGCAATGGCACTGTCTTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 961
DB |||||
4296 TCTTGGGCAATGGCACTGTCTTTGACCAAGCAGAGACTCGGGGGGGGAGACTGACTGTGC 4355
QY 962 TGGCCACCGGTACCCCTCGGGCTCGGTCACTGTGCCCCCATCTTAACATCGAGAGTTG 1021
DB |||||
4356 TGGCCACCGGTACCCCTCGGGCTCGGTCACTGTGCCCCCATCTTAACATCGAGAGTTG 4415
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCCTTTGAAGCAATTA 1081
DB |||||
4416 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCCTTTGAAGCAATTA 4475
QY 1082 AGGGGGGGAGACATCTCTATCTTCTGCCACTCAAAGGAAGTGCAGAGCTCGCGGCAA 1141

```
Db 4476 AGGGGGGAGACATCTCATCTTCTGTCTATTCAAGAAGAGTGCAGCAACTCCGCGAA 4535
Qy 1142 AACTGGTCCGCTTGGCGCTCAATCGCTGGCTTACTACCGCGCTTGTATGTGTCGGTCA 1201
Db 4536 AGCTGGTCCGATTTGGGCATCAATCGCTGGCTTACTACCGCGCTTGTATGTGTCGGTCA 4595
Qy 1202 TCCGCAACAGTGTGTGAGCTGTGCTGCTGGCAACTGAGCCCTCATGACCGGCTTTACCG 1261
Db 4596 TCCGCAACAGCGCGATTTGCTGCTGGTGGCAACCGATGCCCTCATGACCGGCTATACCG 4655
Qy 1362 GCGACTTCGATTCGGTGTAGACTGCACACGCTGTGTACCCAGACATGCATCTACGCC 1321
Db 4656 GCGACTTCGATTCGGTGTAGACTGCATATCGTGTGTACCCAGACATGCATCTACGCC 4715
Qy 1322 TTGACCCCTTACCTTCAACATTTAGACCAATCAGCTTCCCGAGATGCTGTCTCCCGTACTC 1381
Db 4716 TTGACCCCTTACCTTCAACATTTAGACCAATCAGCTTCCCGAGATGCTGTCTCCCGCACTC 4775
Qy 1382 AACGTGCGGGTAGGACTGGCAGAGAGGAGCCAGGATCTACAGATTTGTGGCACCGGGGG 1441
Db 4776 AACGTGCGGGCAGACTGGCAGAGGAGGAGCCAGGATCTACAGATTTGTGGCACCGGGGG 4835
Qy 1442 AGCGTCTTCTGGCATGTTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
Db 4836 AGCGTCTTCTGGCATGTTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4895
Qy 1502 CTTGGTGTAGCTTACGCGCCCGAGACCAAGTGTAGCTTACGAGTGTATGACCGGGTTGTG 1561
Db 4896 CTTGGTGTAGCTTACGCGCCCGAGACTACAGTGTAGCTTACGAGTGTATGACCGGGTTGTG 4955
Qy 1562 CGGACTTCCGCTGTGCCAAGACCATCTTGAATTTTGGAGGGGCTCTTACGGGCTCTCA 1621
Db 4956 CGGACTTCCGCTGTGCCAAGACCATCTTGAATTTTGGAGGGGCTCTTACAGGCTCTCA 5015
Qy 1622 CCCATAGAGCCGCTTCTTATCCCGAGACCAAGAGTGTGGGAAACCTTCCCTATC 1681
Db 5016 CTCATATAGATGCGCACTTCTTATCCCGAGACCAAGAGTGTGGGAAACCTTCCCTATC 5075
Qy 1682 TGGTAGGCTTACAGCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1741
Db 5076 TGGTAGGCTTACAGCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5135
Qy 1742 AGATGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 1801
Db 5136 AGATGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 5195
Qy 1802 ATAGACTGGGCGCTGTCCAGATGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1861
Db 5196 ACAGACTGGGCGCTGTCCAGATGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5255
Qy 1862 TGACATGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1921
Db 5256 TGACATGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5315
Qy 1922 TTCTGGCTGCTTTGGCGCGCTATTCCTGCTATCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1981
Db 5316 TTCTGGCTGCTTTGGCGCGCTATTCCTGCTATCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 5375
Qy 1982 TTGTCTTGTCCGAAAGCCGCAATCATATCCCGACAGGGAAGTCTCTTACCGGAGTTGCG 2041
Db 5376 TCGTCTTGTCCGAAAGCCGCAATCATATCCCGACAGGGAAGTCTCTTACCGGAGTTGCG 5435
Qy 2042 ATGAATGGAAGAGTGTCT 2059
Db 5436 ATGAGATGGAAGAGTGTCT 5453

RESULT 14
AA12710
ID AA12710 standard; cDNA; 9401 BP.
XX
AC AA12710;
```

```
XX 25-MAR-2003 (revised)
DT 15-MAY-1996 (first entry)
XX Hepatitis C virus polyprotein.
XX Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;
KW antibodies; ds.
XX Hepatitis C virus.
XX Key Location/Qualifiers
FT CDS 342..9378
FT /*tag= a
XX EP693687-A1.
XX 24-JAN-1996.
XX 03-APR-1991; 95EP-00114016.
XX 04-APR-1990; 90US-00504352.
XX (CHIR ) CHIRON CORP.
XX Houghton M, Choo Q, Kuo G;
XX WPI; 1996-117956/13.
XX P-PSDB; AAR90931.
XX Combinations of synthetic Hepatitis C Virus antigens - provide more
XX effective diagnosis of Non-A, Non-B Hepatitis.
XX Disclosure; Fig 1(A-Y); 53pp; English.
XX The combination comprises an HCV antigen from the C domain (pref. C22 -
CC AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -
CC AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5
CC (AAR90934) domain. The antigens may in the form of a fusion protein, a
CC simple physical mixture, or the individual antigens commonly bound to a
CC solid matrix. They are pref. prepd. by recombinant DNA techniques
CC (primers are given in AAR12711-12716), but can be synthesised or
CC isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003
CC to correct PF field.)
XX Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;
SQ
Query Match 86.7%; Score 1786; DB 2; Length 9401;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1886; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
Qy 2 TGGCGCTATACGGCCTATGCCAGCAGACACAGGGGCTTTTGGGATGCATATACCA 61
Db 3418 TGGCGCTATACGGCCTATGCCAGCAGACACAGGGGCTTTTGGGATGCATATACCA 3477
Qy 62 GCTTACCGCGCGGACAAACACAGGTGGAGGTGAGGTTCAGATCGTGTCAACTGCTG 121
Db 3478 GCCTACTTGGCGGGACAAACACAGGTGGAGGTGAGGTTCAGATCGTGTCAACTGCTG 3537
Qy 122 CCAGACATTTCTTGGCAACTGTCATTAAACGGGGTGTGTGGACTGTCTACCATGGAGCG 181
Db 3538 CCACAACTTCTTGGCAACTGTCATTAAACGGGGTGTGTGGACTGTCTACCATGGAGCG 3597
Qy 182 GAACAGACCATTGGCTACCTACCAAGGTCTGTATCCAGATGTACACCAATGTGACC 241
Db 3598 GAACAGACCATTGGCTACCTACCAAGGTCTGTATCCAGATGTATACCAATGTGACC 3657
Qy 242 AAGACCTTCGTAGGTGGCGCGCTCCCAAGGTGCCCTCATTAACACCATGTGACTTGG 301
Db 3658 AAGACCTTCGTAGGTGGCGCGCTCCCAAGGTGCCCTCATTAACACCATGTGACTTGG 3717
Qy 302 GCTCCTCGGACCTTTACCTGCTCACAGGACCGCGCATGTCTTCTGTGCGCGACGGG 361
```


Db 3718 GCTCCTCGACCTTTACTCGGTCAAGGACGCGCGATGCTATTCCTCGTGGCGCGGG 3777
Qy 362 GTGATGACAGGGGAGCTGCTTTTCGCCCCGGGCTATCTCTTACTTGAAGGCTCTCTCGG 421
Db 3778 GTGATGACAGGGGAGCTGCTGTCGCCCCGGGCCATTTCTTACTTGAAGGCTCTCTCGG 3837
Qy 422 GAGGCCCTCTGCTGTGTCGCCCGCAGGACATCCGCTAGGCATATTACAGAGCCCGGTATGCA 481
Db 3838 GGGGTCCGCTGTTGTGCCCCCGCGGGCA CGCCGTGGGCAATTTAGGGCCCGGTGTGCA 3897
Qy 482 CCGTGTGAGTGGCTTAAGCCGCTGACTTTCATCCCGCTAGAGAGCTTAGAGCAACCATGA 541
Db 3898 CCGTGTGAGTGGCTTAAGCCGCTGACTTTCATCCCGCTAGAGAGCTTAGAGCAACCATGA 3957
Qy 542 GGTCCCCGGGTGTTCTCAGACAACTCTCCACAGCAGTCCCGCAGAGCTTACCAAGTGG 601
Db 3958 GGTCCCCGGGTGTTCAAGGATACTCTCTCCACAGTAGTGTGCCAGAGCTTCAAGTGG 4017
Qy 602 CCACCTGTGATGTTCCACCGGAGCGGTAAAGACCAAGGTCCCGGCCGATACGCGAG 661
Db 4018 CTCACCTCCATGTTCCACAGGAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCAG 4077
Qy 662 CTCAGGCTACAGAGTCTGCTGCTCAACCCCTCGGTGTGTCGCAACATGGGCTTTGGTG 721
Db 4078 CTCAGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGCAACACTGGGCTTTGGTG 4137
Qy 722 CTACTGACGCCGATCAAGTATTCCACCTACGCGCAAGTTCCTTGGCGAGCGGGTGT 841
Db 4198 CCACCTGACGCCGATCAAGTATTCCACCTACGCGCAAGTTCCTTGGCGAGCGGGTGT 4257
Qy 842 CAGGGGTGCTTATGACATAATTAATTGTGACAGTGCCACTCCAGGATGCAACATCCA 901
Db 4258 CGGGGGCGCTTATGACATAATTAATTGTGACAGTGCCACTCCAGGATGCCACATCCA 4317
Qy 902 TCTTGGSCATGTGCACCTGCTTTCACCAAGCAGAGACCGCGGGGCGAGACTGACGTGTC 961
Db 4318 TCTTGGSCATGTGCACCTGCTTTCACCAAGCAGAGACTGCGGGGGCGAGACTGCTGTGTC 4377
Qy 962 TCGCCACCGCTACCCCTCCGGGCTCCGTCATCTGTGCCCACTCCTTAACATCGAGGAGTTG 1021
Db 4378 TCGCCACCGCACCCCTCCGGGCTCCGTCATCTGTGCCCACTCCTTAACATCGAGGAGTTG 4437
Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAATTA 1081
Db 4438 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAATTA 4497
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGTGGCAGCTCGCGCA 1141
Db 4498 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGTGGCAGCTCGCGCA 4557
Qy 1142 AACTGTGCGGTTGGGCGTCAATGCGGTGCTTACTACCGCGGCTTGTATGTCCGTC 1201
Db 4558 AGCTGTGCGATTGGGCAATCAATGCGGTGCTTACTACCGCGCTTGTAGCTCCGTC 4617
Qy 1202 TCCGACCAAGTGTGAGTGTGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1261
Db 4618 TCCGACCAAGTGTGAGTGTGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 4677
Qy 1262 GGCACCTTCGATTCGGGTGATAGCTGCAACACGCTGTGTCAACAGAGAGTGGCTCAGGC 1321
Db 4678 GGCACCTTCGATTCGGGTGATAGCTGCAACACGCTGTGTGTCAACAGAGAGTGGCTCAGGC 4737
Qy 1322 TTGACCCCTACCTTACCATTTAGACAACTACCGTTCCTCCAGGATGCTGTCTCCCGTACTC 1381
Db 4738 TTGACCCCTACCTTACCATTTAGACAACTACCGTTCCTCCAGGATGCTGTCTCCCGTACTC 4797
Qy 1382 AACCTCGGGTAGGACTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1441
Db 4798 AACCTCGGGTAGGACTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4857

Qy 1442 AGCGTCTTCTGCGCATGTTTGAATCTCGTCTGTCTCTGCGAGTGTCTATGACGCGGTTGTG 1501
Db 4858 AGCGCCCTCTCGGCGATGTTTCGACTCGTCTCGTCTCTGTGAGTGTATGACGAGGCTGTG 4917
Qy 1502 CTTGGTATGAGCTTACGCGCGCGAGACCAACAGTTAGGCTACAGAGCATACATGAACACC 1561
Db 4918 CTTGGTATGAGCTACGCGCGCGAGACTACAGTTAGGCTACAGAGCATACATGAACACC 4977
Qy 1562 CGGAGCTTCCCGTGGCCCAAGACCATCTTGAATTTTGGAGGGCGCTTTACGGGTCTCA 1621
Db 4978 CGGGGCTTCCCGTGGCCAGGACCATCTTGAATTTTGGAGGGCGCTTTACAGGCGCTCA 5037
Qy 1622 CCACATAGACGCGCCACTTCTTATCCAGCAAAAGCAGAGTGGGGAACCTTCCCTATC 1681
Db 5038 CTATATAGATGCCACTTCTTATCCAGCAAAAGCAGAGTGGGGAACCTTCTTACC 5097
Qy 1682 TGGTAGCTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGTCTGTTGGGACC 1741
Db 5098 TGGTAGCTACCAAGCCACCGTGTGCGCTAGGCTCAAGCCCTCCCGTCTGTTGGGACC 5157
Qy 1742 AGATGTGGAAGTCTTGTATCCGCTCAAGCCCACTCCATGGGCCCAACACCTCTGCTAT 1801
Db 5158 AGATGTGGAAGTCTTGTATCCGCTCAAGCCCACTCCATGGGCCCAACACCTCTGCTAT 5217
Qy 1802 ATAGACTGGCGCTGTCCAGAAATGAAGTCACTGACGCAACCCAGTCAACCAATATATCA 1861
Db 5218 ACAGACTGGCGCTGTTCAGAAATGAATCACTGACGCAACCCAGTCAACCAATATATCA 5277
Qy 1862 TGACATGTATGCGGTGACCTGGAGTGTGTCAAGAGTACCTGGGTGTCTGTTGGCGGG 1921
Db 5278 TGACATGTATGCGGTGACCTGGAGTGTGTCAAGAGTACCTGGGTGTCTGTTGGCGGG 5337
Qy 1922 TTCTGGCTCTTGGCGGCTATTCCTATCCAGGCTCGGTGTCTAGTAGTAGGA 1981
Db 5338 TCTGCTGCTTGGCGGCTATTCCTATCCAGGCTCGGTGTCTAGTAGTAGGA 5397
Qy 1982 TTGCTCTTGGCGGCAATCATACCCGACAGGGAAGTCTCTTACCCGGGAGTTG 2041
Db 5398 TCGTCTTGGCGGAGCCGCAATCATACCTGACAGGAGTCTCTTACCCGAGAGTTG 2041
Qy 2042 ATGAATGGAAGTGTCT 2059
Db 5458 ATGAGATGGAAGTGTCT 5475

RESULT 15

AAT99981

ID AAT99981 standard; DNA; 9401 BP.

XX AC AAT99981;

XX AC AAT99981;

XX 25-MAR-2003 (revised)

DT 16-MAR-1998 (first entry)

XX HCV polyprotein coding sequence.

XX HCV polyprotein coding sequence.

XX PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;

XX C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;

XX NS4; ds.

XX Hepatitis C virus.

XX Key

XX Location/Qualifiers

XX 342..9377

XX /*tag= a

XX US5683864-A.

XX 04-NOV-1997.

XX 07-JUL-1992;

XX 92US-00910760.

PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 06-MAY-1988; 88US-00191263.
PR 26-OCT-1988; 88US-00263584.
PR 14-NOV-1988; 88US-00271450.
PR 17-MAR-1989; 89US-00325338.
PR 20-APR-1989; 89US-00341134.
PR 21-APR-1989; 89US-00353896.
PR 18-MAY-1989; 89US-00355002.
PR 04-APR-1990; 90US-00504352.
XX (CHIR) CHIRON CORP.
PA
XX Kuo G, Houghton M, Choo Q;
XX
XX WPI; 1997-548976/50.
DR P-PSDB; AAW34480.
XX
XX Combination of three hepatitis C virus antigens - used for detection of
XX specific antibodies to diagnose infection.
XX
XX Disclosure; Col 25-46; 57pp; English.
XX
CC This sequence represents the Hepatitis C virus polypeptide coding
CC sequence. Fragments of this sequence can be amplified and used in the
CC combination of HCV antigens of the invention. The HCV antigen combination
CC comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)
CC 1-120 of the HCV polypeptide), or its immunologically reactive fragment
CC containing at least 8 aa. It also comprises two additional antigens from
CC two different polypeptide domains, including at least 8 aa from the NS3,
CC NS4, S or NS5 domains of the polypeptide, corresponding, respectively, to
CC aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polypeptide.
CC Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa
CC regions of the HCV polypeptide. These antigen combinations are used
CC diagnostically to detect anti-HCV antibodies, using any standard
CC immunoassay format. These antigen combinations have a broader range of
CC reactivity with antibodies than any antigen individually. (Updated on 25-
CC MAR-2003 to correct PR field.)
XX
XX Sequence 9401 BP; 1883 A; 2860 G; 2673 G; 1985 T; 0 U; 0 Other;
SQ
Query Match 86.7%; Score 1786; DB 2; Length 9401;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
Qy 2 TGGGCGCTATCAGCGCTATGCCAGCAGACAGAGGGGCTTTGGGATGCATATACCA 61
Db 3418 TGGGCGCCATCACGCGCTACGCCAGCAGACAGAGGGGCTCTCTAGGGTGCAATACCA 3477
Qy 62 GCTTACCGCGCGGACAAAACAGGTGGAGGGTGAGGTTGAGTGTCACTGCTG 121
Db 3478 GCCTAATGGCGCGGACAAAACAGGTGGAGGGTGAGGTTGAGTGTCACTGCTG 3537
Qy 122 CCCAGACTTTCTTGGCAACTGCAATTAACGGGGTGTGTGGAGTGTCTACCATGGAGCG 181
Db 3538 CCCAACCCTTCTTGGCAACTGCAATTAACGGGGTGTGTGGAGTGTCTACCATGGAGCG 3597
Qy 182 GAACAGGACCATTTGGTCACTAAGGGTCTGTATCCAGATGTACCAATGTGGACC 241
Db 3598 GAACAGGACCATCGCGTCAACCAAGGGTCTGTATCCAGATGTATACCAATGTAGACC 3657
Qy 242 AAGACCTGTAGGCTGGCGCGCTCCCAAGGTGGCGCTCAATTAACCATGTGCTGCG 301
Db 3658 AAGACCTGTGGCTGGCGCGCTCCGCAAGGTAGCGCGCTCAATTAACCATGTGCTGCG 3717
Qy 302 GCTCTCGGACCTTTACCTGTGTACAGGACGCGCGATGTCAATCTGTGGCGCGAGGG 361
Db 3718 GCTCTCGGACCTTTACCTGTGTACAGGACGCGCGATGTCAATCTGTGGCGCGAGGG 3777
Qy 362 GTGATGGCAGGGGAGCGCTTTTGGCGCGCGCTATCTCTTACTTGAAGGCTCTCTGG 421
Db 3778 GTGATAGCAGGGGAGCGCTGTGTGGCGCGCGCGCTATCTCTTACTTGAAGGCTCTCTGG 3837

Qy 422 GAGGCCCTCTGCTGTGCCCCGCGAGGACATGCGGTAGGCATATTTCAGAGCCCGGTATGCA 481
Db 3838 GGGGTCCGCTGTTGTGCCCCGCGGGGCGAGCGCGTGGGCATATTTTAGGCGCGCGGTGCA 3897
Qy 482 CCGTGGAGTGGCTAAAGCGCGGTGACATTCACCCGCTAGAGAGCTTAGAGACAACCATGA 541
Db 3898 CCGTGGAGTGGCTAAAGCGCGGTGACATTTATCCCTGTGGGAACCTAGAGACAACCATGA 3957
Qy 542 GGTCCCGGTGTTCTCAGACAACCTCCCTCCACAGCAGTGCCTCCAGAGCTTACCAAGTG 601
Db 3958 GGTCCCGGTGTTCCCGATTAATCTCTCTCCACAGTAGTCCCTCCAGAGCTTCCAGGTGG 4017
Qy 602 CCACCTGCTATGCTCCACCGCGAGCGGTAAAGACACCAAGGTCCCGCGCATATCGCAG 661
Db 4018 CTCACTCCATGCTCCACAGGACGCGCAAAAGCAACCAAGTCCCGGCTCATATGCGAG 4077
Qy 662 CTAGGGGCTACAGGTGCTGTGCTCAACCCCTCGTGTGCTGCAACAAATGGGCTTTGGTG 721
Db 4078 CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGCAACACTGGGCTTTGGTG 4137
Qy 722 CTTACATGTCGAAGCCCATGGGATTCATCTTAACATCAGGACTGGGTGAGGACAATTA 781
Db 4138 CTTACATGTCGAAGCTCATGGGATTCATCTTAACATCAGGACCGGGTGAGACAATTA 4197
Qy 782 CTACTGCGAGCCCATCACGTAATTCACCTACGCAAGTTCCTTGGCGGACCGCGGGTGT 841
Db 4198 CCACTGGCAGCCCATCACGTAATTCACCTACGCAAGTTCCTTGGCGGACCGCGGGTGT 4257
Qy 842 CAGGGGTGCTTATGACATTAATTTGTGACAGTGCGCACTCCACGGATGCAACATCA 901
Db 4258 CGGGGGCGCTTATGACATTAATTTGTGACAGTGCGCACTCCACGGATGCAACATCA 4317
Qy 902 TCTTGGCATGGCAGTCTCTTGACCAAGCAGAGACCGGGGGCGGAGACTGCTGTGC 961
Db 4318 TCTTGGCATGGCAGTCTCTTGACCAAGCAGAGACTGCGGGGGCGGAGACTGCTGTGC 4377
Qy 962 TCGCCACCGGTACCCCTCCGGGCTCCGTCACTGTGCCCCCATCTTAAACATCAGAGAGGTG 1021
Db 4378 TCGCCACCGGCACCCCTCCGGGCTCCGTCACTGTGCCCCCATCTTAAACATCAGAGAGGTG 4437
Qy 1022 CTCTGTCCATACCGGAGAGATCCCTTTTATGCAAGGTATTCCTCTTGAAGCAATTA 1081
Db 4438 CTCTGTCCATACCGGAGAGATCCCTTTTATGCAAGGTATTCCTCTTGAAGCAATTA 4497
Qy 1082 AGGGGGGAGACATCTCTCTGCGCACTCAAGAGAGAGTGCAGAGCTCCCGCAA 1141
Db 4498 AGGGGGGAGACATCTCTCTGCGCACTCAAGAGAGAGTGCAGAGCTCCCGCAA 4557
Qy 1142 AACTGTGCGGTGGGGGTCAATGCGCGTGTACTACCGCGGCTTGTATGTGTCGGTCA 1201
Db 4558 AGCTGTGCGATTTGGGCGATCAATGCGCGTGTACTACCGGCTTGTATGTGTCGGTCA 4617
Qy 1202 TCCGACAGTGTGTGAGCTGTGCTGTGCAACTGACGCGCTCATGACCGGCTTTACCG 1261
Db 4618 TCCGACAGTGTGTGAGCTGTGCTGTGCAACTGACGCGCTCATGACCGGCTTTACCG 4677
Qy 1262 GCGACTTTCGATTTGGGCGATCAATGCGCGTGTACTACCGGCTTGTATGTGTCGGTCA 1321
Db 4678 GCGACTTTCGATTTGGGCGATCAATGCGCGTGTACTACCGGCTTGTATGTGTCGGTCA 4737
Qy 1322 TTGACCTCTACTCTTCAACCATTTGAGACAATCAAGTTCCTCCAGGATGTCTCTCCGCTACTC 1381
Db 4738 TTGACCTCTACTCTTCAACCATTTGAGACAATCAAGTTCCTCCAGGATGTCTCTCCGCTACTC 4797
Qy 1382 AAGCTCGGGGTAGGACTGGCAGAGGGGAAGCAGGATCTACAGATTTGTGCAACCGGGG 1441
Db 4798 AAGCTCGGGGTAGGACTGGCAGAGGGGAAGCAGGATCTACAGATTTGTGCAACCGGGG 4857
Qy 1442 AGCGTCTCTTGTGGCATGTTTGAATCGTCTCTCTCGAGTGTCTATGACCGGGGTGTG 1501
Db 4858 AGCGTCTCTTGTGGCATGTTTGAATCGTCTCTCTCGAGTGTCTATGACCGAGGCTGTG 4917

Qy	1502	CTTGGTATAGACTTACGCCCGCGAGACCACAGTTAGGCTACGAGCATACATGAACACC	1561
Db	4918		
Qy	1562	CGGGACTTCCCGTGTGCCAAGACATCTTGAATTTGGGAGGGCGTCTTTACGGGTCTCA	1621
Db	4978		
Qy	1622	CCACATAGACGCCCACTTCCATCCAGACAAAGCAGAGTGGGAAACCTTTCCCTATC	1681
Db	5038		
Qy	1682	TGGTAGCGTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC	1741
Db	5098		
Qy	1742	AGATGTGGAAGTGTTCATCCGCTCAAGCCCACTGACGACCCAGTCACCAAGTATATCA	1801
Db	5158		
Qy	1802	ATAGACTGGCGCTGTCCAGAATGAAGTCACTCAGCAGCCAGTCACCAAGTATATCA	1861
Db	5218		
Qy	1862	TGACATGTATATGCGGCTGACCTGGAGTGTCTACGAGTACCTGGGTGCTCGTTGGCGCG	1921
Db	5278		
Qy	1922	TTCTGGCTGCTTTGGCGCGGTATTCCTATCCACAGGCTCGGTGGTTCATAGTAGGTAGGA	1981
Db	5338		
Qy	1982	TTGCTTTGTCGGAAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTGG	2041
Db	5398		
Qy	2042	ATGAAATGGAAGAGTGCT	2059
Db	5458		

Search completed: February 27, 2005, 18:30:06
Job time : 1064 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 17:06:13 ; Search time 370 Seconds

(without alignments)
9114.504 Million cell updates/sec

Title: US-09-930-591-1

Perfect score: 2061

Sequence: 1 atggcgctatcacggcta.....atgaatggaagagtgtga 2061

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1786	86.7	7310	3	US-08-444-818-74
2	1786	86.7	9379	3	US-09-388-874-1
3	1786	86.7	9379	4	US-09-916-359-1
4	1786	86.7	9401	1	US-07-910-760-9
5	1786	86.7	9401	1	US-08-440-519-9
6	1786	86.7	9401	3	US-08-440-549-9
7	1786	86.7	9401	3	US-08-823-895A-25
8	1785.2	86.6	9401	5	PT-US91-02225-9
9	1784.4	86.6	6785	3	US-08-444-818-65
10	1784.4	86.6	8316	3	US-08-444-818-88
11	1784.4	86.6	8987	3	US-08-444-818-137
12	1784.4	86.6	9185	3	US-08-444-818-122
13	1784.4	86.6	9185	3	US-08-444-818-123
14	1784.4	86.6	9379	3	US-08-444-818-176
15	1781.2	86.4	2058	4	US-09-881-239-2
16	1781.2	86.4	5360	3	US-08-444-818-53
17	1779.6	86.3	2058	4	US-09-881-654-1
18	1779.6	86.3	2058	4	US-10-637-323-1
19	1774.8	86.1	12980	3	US-08-811-566-5
20	1774.8	86.1	12980	3	US-09-034-756-5
21	1773.2	86.0	9646	3	US-08-811-566-1
22	1773.2	86.0	9646	3	US-09-034-756-1
23	1765.2	85.6	9599	3	US-09-014-416-2
24	1765.2	85.6	9599	3	US-09-014-416-6
25	1762	85.5	9416	3	US-08-823-895A-26
26	1762	85.5	9416	4	US-10-104-966-13
27	1750.8	84.9	9401	2	US-08-432-693-1

28	1750.8	84.9	9416	3	US-08-811-566-19
29	1750.8	84.9	9416	3	US-09-034-756-19
30	1749.2	84.9	9365	4	US-09-827-688-7
31	1557.8	75.6	2064	1	US-08-350-884-69
32	1557.8	75.6	2064	1	US-08-440-548-69
33	1557.8	75.6	2064	1	US-08-709-173-69
34	1557.8	75.6	2064	2	US-08-709-177-69
35	1557.8	75.6	2523	1	US-08-350-884-85
36	1557.8	75.6	2523	1	US-08-440-548-85
37	1557.8	75.6	2523	1	US-08-709-173-85
38	1557.8	75.6	2523	2	US-08-709-177-85
39	1546	75.0	1932	3	US-09-128-314-1
40	1546	75.0	8157	3	US-09-128-314-3
41	1398.4	67.9	9413	4	US-09-827-688-6
42	1381.2	67.0	7989	4	US-09-539-601-10
43	1381.2	67.0	8001	4	US-09-539-601-7
44	1379.6	66.9	8001	4	US-09-539-601-22
45	1379.2	66.9	8638	4	US-10-029-907-7

ALIGNMENTS

RESULT 1
US-08-444-818-74
; Sequence 74, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..7310
US-08-444-818-74

Query Match 86.7%; Score 1786; DB 3; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1886; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGCGGCTATACGGCTATGCCCCAGACAGAGAGGGGCGCTTTTGGGATGCAATACCA 61
Db 1729 TGGCGCCCATACGGGTCAGCCAGAGAGCAAGAGGGCCCTCCTAGGGTCATATACCA 1788
QY 62 GCTTGACCGGCGGGCAAAACACAGGTGGAGGTGAGGTTCAGATCGTGTCAACTGCTG 121
Db 1789 GCCTAACTGGCGGGCAAAACCAAGTGGAGGTGAGGTTCAGATGTTGTCAACTGCTG 1848
QY 122 CCAGACTTTCCTGGCAACCTGATTAACGGGGTGTGGAGTGTCTACCATGGAGCGG 181
Db 1849 CCCAAACCTTCTGGCAACCTGATCAATGGGGTGTGGAGTGTCTACCATGGAGCGG 1908
QY 182 GAACAGGACATTTGGTCACTTAAGGTCCTCTATCCAGATGTACACCAATGTGGACC 241
Db 1909 GAACAGGACATCGGTCACCAAGGTCTCTGTATCCAGATGTATCCAGATGTATCCAGATGTAGACC 1968
QY 242 AAGACCTCTGAGGCTGGCGGCTCCCAAGGTGCGGCTCATTAACACCATGCACTTGG 301
Db 1969 AAGACCTTGTGGCTGGCGGCTCCCAAGGTAGCGCTCATTAACACCATGCACTTGG 2028
QY 302 GCTCCTCGACCTTTACCTGGTCACAGGACACCGCATGTCTCTGTCGGCGGACGG 361
Db 2029 GCTCCTCGACCTTTACCTGGTCACAGGACACCGCATGTCTCTGTCGGCGGACGG 2088
QY 362 GTGATGGCAGGGCAGCTGCTTTGCGCCCGGCTATCTCTTACTTTGAAAGGCTCCTCGG 421
Db 2089 GTGATGAGGGGAGCTGCTGTCGCGCCCGGCTATCTCTTACTTTGAAAGGCTCCTCGG 2148
QY 422 GAGGCGCTCTGCTGTCGCGCGAGGACATGCGGTAGGCATATTTCAGAGCGCGGTATGCA 481
Db 2149 GGGGTCGCTGTTGTGCGCGGGGCAAGCGGTGGGCATATTTCAGGCGCGGTGTGCA 2208
QY 482 CCGGTGAGGTGCTAAGCGGTGGAATTCATCCCGGTAGAGCTTTAGAGACCAATGA 541
Db 2209 CCGGTGAGGTGCTAAGCGGTGGAATTCATCCCGGTAGAGCTTTAGAGACCAATGA 2268
QY 542 GGTCCCGGTGTTCTCAGACCACTCTCCCGGACAGGTCGCCAGAGTACCAAGTGG 601
Db 2269 GGTCCCGGTGTTCTCAGACCACTCTCCCGGACAGGTCGCCAGAGTACCAAGTGG 2328
QY 602 CCACCTGCACTGCTCCCGGAGGCGGTAAAGACCAAGGTCGCCCGGCGCATACGAG 661
Db 2329 CTCACTCCATGCTCCCGGAGGCGGCAAAAGCACCAAGGTCGCCCGGCGCATACGAG 2388
QY 662 CTAGGGCTACAGGTGCTGCTGCTCAACCCCTCGGTGCTGCAACAAATGGGCTTTGGTG 721
Db 2389 CTAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACAACTGGGCTTTGGTG 2448
QY 722 CTTTACATGTCAGGGCCATGGGATTCATCTAACATCAGGACTGGGGTGAGGACAAATTA 781
Db 2449 CTTTACATGTCAGGGCTCATGGATCGATCTAACATCAGGACCGGGTGAGGACAAATTA 2508
QY 782 CTACTGACGCGCGATCAGATATTCACCTACGGCAAGTTCCTTTGCGAGCGGGGTGTT 841
Db 2509 CCACTGSCAGCCCATCACTGCTCCACCTACGGCAAGTTCCTTTGCGAGCGGGGTGCT 2568
QY 842 CAGGGGTGCTTATGACATTAATTTGTACAGGTGCCATCCACGGATGCAACATCCA 901
Db 2569 CAGGGGTGCTTATGACATTAATTTGTACAGGTGCCATCCACGGATGCAACATCCA 2628
QY 902 TCTTGGGCATTTGACATGCTCTTGACCAAGCAGAGACCGGGGGGCGAGACTGACATGTGC 961
Db 2629 TCTTGGGCATTTGACATGCTCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTGC 2688
QY 962 TCGCCACCGGTACCCCTCCGGGCTCGGTCACTGTGCGCCCATCTTAACATCAGAGGTTG 1021
Db 2689 TCGCCACCGGCACCCCTCCGGGCTCGGTCACTGTGCGCCCATCTTAACATCAGAGGTTG 2748
QY 1022 CTCTGTCCACTACCGGAGATCCCTTTTATGCAAGGCTATCCCTTTGAGGCAATTA 1081
Db 2749 CTCTGTCCACTACCGGAGATCCCTTTTATGCAAGGCTATCCCTTTGAGGCAATTA 2808

QY 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAAAGTGCAGACGCTGCGCGCAA 1141
Db 2809 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAAAGTGCAGACGCTGCGCGCAA 2868
QY 1142 AACTGGTCCGCTTGGCGTCAATGCGGTGCTTACTACCGCGGCTTGTGATGTGCTGCTCA 1201
Db 2869 AGCTGGTCCGCTTGGCGTCAATGCGGTGCTTACTACCGCGGCTTGTGATGTGCTGCTCA 2928
QY 1202 TCCGACACAGTGTGACGTTGTGCTGCGCAACTGACGCCCTCATGACCGGCTTTACCG 1261
Db 2929 TCCGACACAGCGGCTGATGTTGTGCTGCGCAACCGATGCTCATGACCGGCTTACCG 2988
QY 1262 GCGACTTCGATTCGGGTAGACTGCAACACGTTGTGTACCCAGACAGTGCAGCTTCAGCC 1321
Db 2989 GCGACTTCGACTCGGTGATAGACTGCAATACGTTGTGTACCCAGACAGTGCAGCTTCAGCC 3048
QY 1322 TTGACCTTACCTTACCATTTGACATCAACGTTTCCCGAGGATGCTGCTCCCGTACTC 1381
Db 3049 TTGACCTTACCTTACCATTTGACATCAACGTTTCCCGAGGATGCTGCTCCCGTACTC 3108
QY 1382 AACGTCGGGCTAGGACTGCGCAGAGGAAAGCCAGGCACTTACAGATTTGTGGCACCGGGG 1441
Db 3109 AACGTCGGGCTAGGACTGCGCAGAGGAAAGCCAGGCACTTACAGATTTGTGGCACCGGGG 3168
QY 1442 AGCGTCCCTTCTGGCATGTTTGAATGCTGCTCTGCGAGTGTATGACCGGGGTGTG 1501
Db 3169 AGCGTCCCTTCTGGCATGTTTGAATGCTGCTCTGCGAGTGTATGACCGAGGCTGTG 3228
QY 1502 CTTGTTATGAGCTTACCGCGCGGAGACACAGTTAGGCTAGGACATACATGAACACCC 1561
Db 3229 CTTGTTATGAGCTTACCGCGCGGAGACACAGTTAGGCTAGGACATACATGAACACCC 3288
QY 1562 CGGGACTTCCCGTGTGCCAAGACCATCTTTGAATTTTGGAGGCGCTTTTACGGGCTCTCA 1621
Db 3289 CGGGGCTTCCCGTGTGCCAAGACCATCTTTGAATTTTGGAGGCGCTTTTACAGGCTCTCA 3348
QY 1622 CCCACATAGACGCCCTTCTTATCCAGACAAAGCAGAGTGGGGAAAACCTTTCCTATC 1681
Db 3349 CTTATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTTCCTTACC 3408
QY 1682 TGGTAGCTACCAAGCCAGGCTGCGGTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
Db 3409 TGGTAGCTACCAAGCCAGGCTGCGGTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 3468
QY 1742 AGATGTGAAGTGTCTTGTATCCGCTCTCAAGCCACCTCCATGGGCGCAACACCTCTGCTAT 1801
Db 3469 AGATGTGAAGTGTCTTGTATCCGCTCTCAAGCCACCTCCATGGGCGCAACACCTCTGCTAT 3528
QY 1802 ATAGACTGGGCGCTGTCCAGATGAAGTCAACCTGTGACCGACCCAGTCAACAGTATATCA 1861
Db 3529 ACAGACTGGGCGCTGTTCAGAAATGAATCAACCTGTGACCGACCCAGTCAACAAATATATCA 3588
QY 1862 TGACATGATGTGCGCTGACCTGGAGTGTGTACAGTACCTGGGTGCTGTTGGGCGG 1921
Db 3589 TGACATGATGTGCGCTGACCTGGAGTGTGTACAGTACCTGGGTGCTGTTGGGCGG 3648
QY 1922 TTTCTGCTGCTTTGGCGGCTATTGCTTATCCAGGCTGCTGCTGCTATAGTAGTAGGAGGA 1981
Db 3649 TTTCTGCTGCTTTGGCGGCTATTGCTTATCCAGGCTGCTGCTGCTATAGTAGTAGGAGGA 3708
QY 1982 TTGTCCTTGTCCGAAAGCGCGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTG 2041
Db 3709 TCGTCTTGTCCGGAAGCGCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTG 3768
QY 2042 ATGAATGGAAGTGTCT 2059
Db 3769 ATGAGATGGAAGTGTCT 3786

RESULT 2
US-09-388-874-1
; Sequence 1, Application US/09388874
; Patent No. 6284249

GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; TITLE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: PMCF97-03A
; CURRENT APPLICATION NUMBER: US/09/388,874
; CURRENT FILING DATE: 1993-09-02
; EARLIER APPLICATION NUMBER: PCT/FR98/00448
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 97/02,887
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320)...(9352)
us-09-388-874-1

Query Match 86.7%; Score 1786; DB 3; Length 9379;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy	2	TGGCGCCTATCACGGCCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCAATATCACCA	61
Db	3396	TGGCGCCCATCAAGCGGTACGCCAGCAGACAAAGGGGCTCTTAGGGTGCATATCACCA	3455
Qy	62	GCTTGACGGCGGGACAAACACAGGTGGAGGTGAGGTTGAGATCGTGTCAACTGCTG	121
Db	3456	GCCTAACTTGGCGGGACAAACACAGGTGGAGGTGAGGTTGAGATCGTGTCAACTGCTG	3515
Qy	122	CCAGAGCTTCTTGGCAACCTGCATTAACGGGCTGTTGGAGCTGTACCATCGAGCGG	181
Db	3516	CCCAACCTTCTTGGCAACCTGCATTAATGGGCTGTTGGAGCTGTACCATCGAGCGG	3575
Qy	182	GAAACAAGGACCATTCGCTACCTAAGGCTCCTGTTATCCAGATGATACCAATGTGGACC	241
Db	3576	GAACGAGGACCATTCGCTACCTAAGGCTCCTGTCATCCAGATGATACCAATGTAGACC	3635
Qy	242	AAGACCTGTAGGTGGCGGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGGG	301
Db	3636	AAGACCTTGTGGCTGGCGGCTCCCAAGGTAGCGCTCATTAACACCATGCACTTGGG	3695
Qy	302	GCTCCTCGGACCTTACTGTGTACAGGACGCGGATGTCATTCCTGTGGCGGACGGG	361
Db	3696	GCTCCTCGGACCTTACTGTGTACAGGACGCGGATGTCATTCCTGTGGCGGACGGG	3755
Qy	362	GTGATGGCAGGGCAGGCTGCTTTTCGCCCGGCTTCTTACTTGAAGGCTCCTCGG	421
Db	3756	GTGATAGCAGGGCAGGCTGCTGTGCGCCCGGCCATTTCTTACTTGAAGGCTCCTCGG	3815
Qy	422	GAGGCCCTTGTCTGTGCGCGCAGGACATGCGGTAGGCATATTCAGAGCGCGGTATGCA	481
Db	3816	GGGTCGCGCTGTTGTGCGCGCGGACGCGGTGGGCATATTTAGGGCGCGGTGTGCA	3875
Qy	482	CCCGTGGAGTGGCTTAAGCGGTGACCTTCATCCCGGTAGAGAGCTTAGAGACACCATGA	541
Db	3876	CCCGTGGAGTGGCTTAAGCGGTGACCTTCATCCCGGTAGAGAGCTTAGAGACACCATGA	3935
Qy	542	GGTCCCGGCTTCTCAGACAACTCCTCCCAACAGCAGATGCCCGCAGAGCTTACCAAGTGG	601
Db	3936	GGTCCCGGCTTCTCAGGATTAACCTCCTCCACAGTAGTGTCCCGCAGAGCTTCCAGGTGG	3995
Qy	602	CCCACTGTGCTATCCCAACGGCAGCGGTAAAGACCAAGGTCCCGCGCGCATACGAG	661
Db	3996	CTCACTCCATGCTCCCAACGGCAGCGGTAAAGACCAAGGTCCCGCGCTGATATGAG	4055
Qy	662	CTCAGGCTACAAGGTGCTGTCTCAACCCCTCGGTGCTGCAACATGGGGCTTTGGTG	721
Db	4056	CTCAGGCTACAAGGTGCTGTCTCAACCCCTCGGTGCTGCAACATGGGGCTTTGGTG	4115

Qy	722	CTTACATGTCCAAGGCCCATGGGATTGATCTTAACATCAGAGCTGGGTGAGGCAATTA	781
Db	4116	CTTACATGTCCAAGGCTCATGGGATCGATCTTAACATCAGAGCGGGTGAGAACAAATTA	4175
Qy	782	CTACTGSCAGCCCGATCAGCTATTCACCTACGGCAAGTTCTTGGCCGACGGCGGGTGT	841
Db	4176	CCACTGGCAGCCCGATCAGCTATTCACCTACGGCAAGTTCTTGGCCGACGGCGGGTGT	4235
Qy	842	CAGGGGTGTATGACATAATATTTGTGACGAGTGCCACTCCACGATGCAACATCCA	901
Db	4236	CGGGGGCGCTTATGACATAATATTTGTGACGAGTGCCACTCCACGATGCAACATCCA	4295
Qy	902	TCTTGGGCATTTGGCATTCTCTTGAACCAAGCAGAGACCGGGGGGCGAGACTGACTGTC	961
Db	4296	TCTTGGGCATTTGGCATTCTCTTGAACCAAGCAGAGACTGCGGGGGGCGAGACTGCTGTG	4355
Qy	962	TGCGCACCGCTACCCCTCCGGGCTCGGTCACTGTGCCCCCATCTTAACATCAGAGAGTTG	1021
Db	4356	TCGCCACCGCACCCCTCCGGGCTCGGTCACTGTGCCCCCATCTTAACATCAGAGAGTTG	4415
Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTTGAAGCAATTA	1081
Db	4416	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTTGAAGCAATTA	4475
Qy	1082	AGGGGGGAGACATCTCATCTTCTGCACTCAAGAAAGAGTGCAGAGCTCCGCCAA	1141
Db	4476	AGGGGGGAGACATCTCATCTTCTGCACTCAAGAAAGAGTGCAGAGCTCCGCCAA	4535
Qy	1142	AACGTGTCGGTGGGGCTCAATGCCGTGCTTACTACCGGGCTTGTATGTGCTGCTCA	1201
Db	4536	AGCTGTCGATTTGGGCATCAATGCCGTGCTTACTACCGGGCTTGTATGTGCTGCTCA	4595
Qy	1202	TCCGACACGAGTGTGAGTTGTGCTGTGCACTGACGCCCTCATGACCGGCTTTTACCG	1261
Db	4596	TCCGACACGAGGCTGTTGTGCTGTGCACTGACGCCCTCATGACCGGCTTTTACCG	4655
Qy	1262	GCACATTTCTGATGATAGACTGCAACACGTTGTGCTCACCAGACAGTGCAGCTTCAGCC	1321
Db	4656	GCACATTTCTGATGATAGACTGCAACACGTTGTGCTCACCAGACAGTGCAGCTTCAGCC	4715
Qy	1322	TTGACCTTACTCTTACCATTTGAGACATCAGCTTCCCGAGGATGCTGTCTCCCGTACTC	1381
Db	4716	TTGACCTTACTCTTACCATTTGAGACATCAGCTTCCCGAGGATGCTGTCTCCCGTACTC	4775
Qy	1382	AACGTGCGGGTAGGACTGCGCAGAGGGAAGCAGCATCTACAGATTTGTGSCACCGGGG	1441
Db	4776	AACGTGCGGGTAGGACTGCGCAGAGGGAAGCAGCATCTACAGATTTGTGSCACCGGGG	4835
Qy	1442	AGCGTCCCTTCTGGCATGTTTGAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1501
Db	4836	AGCGTCCCTTCTGGCATGTTTGAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	4895
Qy	1502	CTTGTGATGAGCTTACGCGCGCGAGACACAGTGTAGGCTACGAGCATACATGAACACCC	1561
Db	4896	CTTGTGATGAGCTTACGCGCGCGAGACACAGTGTAGGCTACGAGCATACATGAACACCC	4955
Qy	1562	CGGACTTCCCGTGTGCCAAGACCATCTGAAATTTTGGGAGGGGCTCTTTACGGGTCTCA	1621
Db	4956	CGGACTTCCCGTGTGCCAAGACCATCTGAAATTTTGGGAGGGGCTCTTTACGGGTCTCA	5015
Qy	1622	CCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGGGAAACCTTCCCTATC	1681
Db	5016	CTCATATAGATGCCACTTCTTATCCAGACAAAGCAGAGTGGGGGAAACCTTCTTACC	5075
Qy	1682	TGTTAGGCTACCAAGCACCAGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGCTGGGACC	1741
Db	5076	TGTTAGGCTACCAAGCACCAGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGCTGGGACC	5135
Qy	1742	AGATGTGAAGTGTGATCCGCTCAAGCCCAACCTCCATGGGCCAACACCTCTGCTAT	1801
Db	5136	AGATGTGAAGTGTGATCCGCTCAAGCCCAACCTCCATGGGCCAACACCTCTGCTAT	5195

QY	1802	ATAGACTGGGCGCTGTGCAGAAATGAAGTCAACCTTGACGACACCCAGTGCACCAAGTATATCA	1861
Db	5196	ACAGACTGGGGCGCTGTTTCAGAAATGAAATCACCTTGACGACACCCAGTGCACCAAAATACATCA	5255
QY	1862	TGACATGTATCTCGGCTGACCTCGAGGTGCTGACGAGTACCTGGGTGCTCGTTGGCGGCG	1921
Db	5256	TGACATGCATGTTCGGCGACCTGGAGGTGCTCACGAGCACTGGGTGCTGTTGGCGGCG	5315
QY	1922	TTCCTGGCTGCTTTTGGCCGCGTATTTGCCCTATCCACAGGCTCGTGGTTCATAGTAGGTAGGA	1981
Db	5316	TCCTGGCTGCTTTTGGCCGCGTATTTGCCCTGTCAAAGCTGCTGATAGTGGCAGGG	5375
QY	1982	TTGCTTTGTCGGGAAGCCGCAATCATACCCGACAGGAGTCTCTACCGGAGTTTCG	2041
Db	5376	TCGCTTTGTCGGGAAGCCGCAATCATACCTTGACAGGAAGTCTCTACCGAGAGTTTCG	5435
QY	2042	ATGAAATGGAAGAGTGCT	2059
Db	5436	ATGAGATGGAAGAGTGCT	5453
RESULT 3			
US-09-916-359-1			
; Sequence 1, Application US/09916359			
; Patent No. 6538123			
; GENERAL INFORMATION:			
; APPLICANT: Veronique Barban			
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR			
; TITLE OF INVENTION: TREATING C HEPATITIS			
; FILE REFERENCE: PMC97-03A			
; CURRENT APPLICATION NUMBER: US/09/916,359			
; CURRENT FILING DATE: 2001-07-26			
; PRIOR APPLICATION NUMBER: 09/388,874			
; PRIOR FILING DATE: 1999-09-02			
; PRIOR APPLICATION NUMBER: 97/02,887			
; PRIOR FILING DATE: 1997-03-06			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 1			
; LENGTH: 9379			
; TYPE: DNA			
; ORGANISM: Virus			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (320)...(9352)			
US-09-916-359-1			
Query Match			
Best Local Similarity 86.7%; Score 1786; DB 4; Length 9379;			
Matches 188; Conservative 0; Mismatches 170; Indels 0; Gaps 0;			
QY	2	TGGCGCTATCAGGCGCTATGCCAGCAGACAAGGGCGCTTTGGGATGCATAATCACCA	61
Db	3396	TGGCGCCATCAGGCGTACGCCAGCAGACAAGGGCGCTCTTAGGTGCATATCACCA	3455
QY	62	GCTTGACCGGCGGGGACAAAACACAGGTGAGGGTGAGGTTTCAGATCGTGTCAACTGCTG	121
Db	3456	GCCTAACTGCCGGGACAAAACCAAGTGAGGGTGAGGTCCAGATTGTTCAACTGCTG	3515
QY	122	CCACAGCTTTTGGCAACTGCATTAAACGGGGTGTTGGACTGTCTACCATGGAGCGG	181
Db	3516	CCCAACCTTCTGGCAACGTGCATCAATGGGGTGCTGGACTGTCTACCGGGCGG	3575
QY	182	GACCAAGGACATTCGTGTCACCTTAAGGGTCTCTTTATCCAGATGTATACCAATGTGGACC	241
Db	3576	GACCGAGGACATTCGTGTCACCAAGGGTCTCTGTCATCCAGATGTATACCAATGTAGACC	3635
QY	242	AAGACCTCGTAGGTGGCCCGCTCCCAAGGTCCCGCTCATTTAACACCATGCACTTGG	301
Db	3636	AAGACCTTGTGGCTGGCCCGCTCCGCAAGGTAGCGCTCAITTGACACCTTGCACTTGG	3695
QY	302	GCTCTCTGGACCTTTTACCTGGTCACGAGGCACGCCGATGTCAATTCTGTGGCGCGACGG	361


```
QY 1442 AGCGTCCCTTGGGCAAGTTTGAATGCTGCTGTCCTCTGGAGTGTATGACGCGGTTGTG 1501
| | | | |
Db 4836 AGCGCCCTCCGGCATGTTTCGACTCGTCCGTCCTCTGTGAGTGTATGACGCGGCTGTG 4895
| | | | |
QY 1502 CTTGTATGAGTTACGCGCCGCGAGACACAGTTTAGCTAGGACATACATGACACCC 1561
| | | | |
Db 4896 CTTGTATGAGTTACGCGCCGCGAGACTACAGTTAGCTAGGACGATACATGACACCC 4955
| | | | |
QY 1562 CGGACTTCCCGTGTGCCAAGACCACTTTGAATTTGGAGGGCTCTTTACGGGTCTCA 1621
| | | | |
Db 4956 CGGGCTTCCGCTGTGCCAGGACCATCTTGAATTTGGAGGGCTCTTTACAGGCTCA 5015
| | | | |
QY 1622 CCCATAGAGCCGCTCTCTCTATCCAGACAAAGACAGTGGGAAACCTTCCCTATC 1681
| | | | |
Db 5016 CTCATATAGATGCCCACTTCTATCCAGACAAAGACAGTGGGAGAACCTTCTTACC 5075
| | | | |
QY 1682 TGGTAGCTTACCAAGCCAGCCGCTGCGCTAGAGCTCAAGCCCTCCCGCTGTTGGACC 1741
| | | | |
Db 5076 TGGTAGCTTACCAAGCCAGCCGCTGCGCTAGAGCTCAAGCCCTCCCGCTGTTGGACC 5135
| | | | |
QY 1742 AGATGTGAAGTGTGATCGCTCTCAAGCCACCCCTCCATGGGCCAACACCTCTGCTAT 1801
| | | | |
Db 5136 AGATGTGAAGTGTGATTTGGCTCAAGCCACCCCTCCATGGGCCAACACCCCTGCTAT 5195
| | | | |
QY 1802 ATAGACTGGGCGTGTCCAGATGAAGTCAACCTGACGCCACCCAGTCAACCAAGTATATCA 1861
| | | | |
Db 5196 ACAGACTGGGCGTGTTCAGAATGAATCAACCTGACGCCACCCAGTCAACCAATATATCA 5255
| | | | |
QY 1862 TGACATGATGTCGGTGACCTGGAGTGTCTCAGAGTACCTGGGTGCTGTTGGCGCG 1921
| | | | |
Db 5256 TGACATGATGTCGGCGCCGACCTGGAGTGTCTCAGAGCACCTGGGTGCTGTTGGCGCG 5315
| | | | |
QY 1922 TTCTGGCTGCTTTGGCCGCTATTGCCCTATCCACAGGCTGCTGATAGTAGTAGGA 1981
| | | | |
Db 5316 TCCTGGCTGCTTTGGCCGCTATTGCCCTATCCACAGGCTGCTGATAGTAGTAGGA 5375
| | | | |
QY 1982 TTGTCTTTCGGAAGCGGCAATCATACCCGACAGGAAAGTCTCTACCGGAGTTTCG 2041
| | | | |
Db 5376 TCGTCTTTCGGAAGCGGCAATCATACCTGACAGGAAAGTCTCTACCGAGAGTTTCG 5435
| | | | |
QY 2042 ATGAATGGAAGTGCT 2059
| | | | |
Db 5436 ATGATGGAAGTGCT 5453
| | | | |

RESULT 4
; Sequence 9, Application US/07910760
; Patent No. 5683864
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07910,760
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9401 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 342..9374
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 366
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Arg."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 372
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 867
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1341
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2148
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ile."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2883
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Asn."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3681
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ser."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3690
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4167
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Leu."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4323
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4701
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Tyr."
; FEATURE:
```

```
/ NAME/KEY: misc_feature
/ LOCATION: 4752
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 5970
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6183
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be His."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6186
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Cys."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6402
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7386
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7494
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7497
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ala."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7845
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 8409
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9102
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9327
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Pro."
/ US-07-910-760-9
Query Match 86.7%; Score 1786; DB 1; Length 9401;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 2 TGGCGCTATCAGCGCTATGCCAGCAGAAAGGGCCCTTTGGGATGCAATATCACCA 61
DB 3418 TGGCGCCATCAGCGGTACGCCGTCAGCAGAAAGGGCCCTCTAGGGTGCAATACCA 3477
QY 62 GCTTGACCGCGGGACAAAACCCAGGTGAGGTTCAGATCGTGTCAACTGCTG 121
DB 3478 GCCTAACTGCGCGGGACAAAACCAAGTGAGGGTGAGGCCAGATTGTGTCAACTGCTG 3537
122 CCAGACTTTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCCG 181
3538 CCCAAACCTTCTCTGGCAAGTGCATCAATGGGGTGTGTGGACTGTCTACCATGGAGCCG 3597
182 GAACAGGACCATTTGGTCACTTAAGGTCTCTGTATCCAGATGTATACCAATGTGAC 241
3598 GAACGAGGACCATCGCGTCACCAAGGGTCTGTCTATCCAGATGTATACCAATGTGAC 3657
242 AAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCCGCTCATTTAAACCATGCACTTGG 301
3658 AAGACCTTGTGGGTGGCCCGCTCCGCAAGGTAGCGCTCATTTGACACCTGCACTTGG 3717
302 GCTCTCGGACCTTTACCTTGTGTCAGAGGACCGCGATGTCACTTCTGTGGCGGACGG 361
3718 GCTCTCGGACCTTTACCTTGTGTCAGAGGACCGCGATGTCACTTCTGTGGCGGACGG 3777
362 GTGATGGCAGGGGACGCTCTTTCGCCCGGCTATCTTACTTTGAAGGCTCTCTCG 421
3778 GTGATAGCAGGGGACGCTCTTTCGCCCGGCTATCTTACTTTGAAGGCTCTCTCG 3837
422 GAGGCCCTCTGTGTGCCCCCGAGGACATGCGGTAGGCATATTCAGAGCGCGGTATGCA 481
3838 GGGTCCGCTGTGTGCCCCCGGGGACGCGGTGGCATATTTAGGGCGCGGTGTGCA 3897
482 CCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGGTAGAGACTTTAGAGCAACCATGA 541
3898 CCGTGGAGTGGCTAAGCGGTGGACTTCATCCCTGGGAACTTAGAGCAACCATGA 3957
542 GGTCCCCGGTGTCTCAGACAACTCTCCCAAGGAGTGGCCAGAGCTACCAAGTGG 601
3958 GGTCCCCGGTGTCTCAGGATAACTCTCTCCAGTAGTGGCCCCAGAGCTTCAGGTGG 4017
602 CCACCTGCATGCTCCACCGGAGCGGTAAAGACCAAGGTCCCGGCGCATATAGCCAG 661
4018 CTCACCTGCATGCTCCACCGGAGCGGTAAAGACCAAGGTCCCGGCGCATATAGCCAG 4077
662 CTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTGTGCAACAATGGGGTGGTG 721
4078 CTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTGTGCTGCAACAATGGGGTGGTG 4137
722 CTTCATGTCCAAAGGCCCATGGATGTATCTTAACATCAGGACTGGGGTGAGGACAATTA 781
4138 CTTCATGTCCAAAGGCCCATGGATGTATCTTAACATCAGGACTGGGGTGAGGACAATTA 4197
782 CTACTGGCAGCCGATCAGTATTCACCTCAGGCAAGTTCCTTGGCGAGCGGGTGT 841
4198 CCACTGGCAGCCGATCAGTATTCACCTCAGGCAAGTTCCTTGGCGAGCGGGTGT 4257
842 CAGGGGTGCTTATGACATAAATTTGTGACGAGTGCCACTCCACGATGCAACATCCA 901
4258 CGGGGGGCGTTATGACATAAATTTGTGACGAGTGCCACTCCACGATGCAACATCCA 4317
902 TCTTGGGCAATGGCACTGTCTTGACCAAGCAGAGACCGGGGGGCGAGACTGACTGTC 961
4318 TCTTGGGCAATGGCACTGTCTTGACCAAGCAGAGACTGCGGGGGGCGAGACTGCTGTGTC 4377
962 TCGCCACCGCTACCCCTCGGGGTCCGTCACTGTGCCCATCTTAACATCAGGAGGTTG 1021
4378 TCGCCACCGCCACCCCTCGGGGTCCGTCACTGTGCCCATCTTAACATCAGGAGGTTG 4437
1022 CTCGTGCTCACTACCGGAGATCCCTTTATGCGCAAGGCTATTCCTTTGAAGCAATTA 1081
4438 CTCGTGCTCACTACCGGAGATCCCTTTATGCGCAAGGCTATTCCTTTGAAGCAATTA 4497
1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGGCGAGAGTCCGCGCAA 1141
4498 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGGCGAGAGTCCGCGCAA 4557
1142 AACTGTGCGGTGGCGGTCAATGCGGTGGCTTACTACCGCGGCTTGTGTGCTGCTCA 1201
4558 AGTGTGCGATTTGGGATCAATGCGGTGGCTTACTACCGCGGCTTGTGTGCTGCTCA 4617
1202 TCCCGACCAAGTGTGAGTGTGCTGCTGGGCAACTGAGCCCTCATGACCGGCTTTACCG 1261
```



```
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4167
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Leu."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4323
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4701
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Tyr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4752
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 5970
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6183
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be His."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6186
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Cys."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6402
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7386
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7494
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7497
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ala."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7845
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 8409
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9102
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also by Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
```

```
/ LOCATION: 9327
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Pro."
/ US-08-440-519-9

Query Match      86.7%; Score 1786; DB 1; Length 9401;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY      2  TGGCGCCTATCAGCGCCTATGCCAGACAGACAGAGGGCCCTTTTGGGATGCATATACCA 61
DB      3418 TGGCGCCCATCAGCGGTAGCGCCAGACACAAAGGGCCCTCTAGGGTGCAATATACCA 3477

QY      62  GCTTGACCGCGCGGACAAAACCAAGTGGAGGGTGAGGTTCAAGATCGTGTCAACTGCTG 121
DB      3478 GCCTAACTGGCGGGACAAAACCAAGTGGAGGGTGAGGTTCCAGATGTTGTCAACTGCTG 3537

QY      122 CCCAGACTTTTGGCAACTGCTGCAATTAACGGGGTGTGTTGGAGTCTCTACCATGGAGCCG 181
DB      3538 CCCAAACCTTCTGTGGCAACGTGCATCAATGGGGTGTGCTGGAGTGTCTACACGGGGCCG 3597

QY      182 GAACAGGACCATTGGCTCACCTAAGGTCCTGTTATCCAGATGTACACCAATGTGGACC 241
DB      3598 GAACGAGGACCATCGCGTCACCAAGGGTCCTGTCAATCCAGATGTATACCAATGTGAGCC 3657

QY      242 AAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGGC 301
DB      3658 AAGACCTTGTGGCTGGCGCGCTCCCAAGGTAGCGCTCATTTGACACCTGCACTTGGC 3717

QY      302 GCTCTCGAACCTTTACTCTGGTCAAGAGCAGCGCGATGTCATTCCTGTCGCGCGGCGG 361
DB      3718 GCTCTCGGACCTTTACTTGGTCAAGAGCAGCGCGATGTCATTCCTGTCGCGCGGCGG 3777

QY      362 GTGATCGCAGGGGCGAGCTGCTTTCCCGCGGCTATCTTACTTTGAAGGCTCTCCG 421
DB      3778 GTGATAGCAGGGGCGAGCTGCTGTCCCGCGGCGCAATTCCTACTTTGAAAGGCTCTCCG 3837

QY      422 GAGGCCCTCTGTGTGCCCGCAGAGACATGCCGTAGGCATATTCAGAGCGCGCGGTATGCA 481
DB      3838 GGGGTCCGCTGTTGTGCCCGCGGGGACGCGGTGGGCATATTTAGGGCGCGCGGTGCA 3897

QY      482 CCGGTGAGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGACTTAGAGAACCAATGA 541
DB      3898 CCGGTGAGTGGCTAAGCGGTGGACTTTCATCCCTGTGGAGAACCTAGAGAACCAATGA 3957

QY      542 GGTCCCGGGTGTCTCAGACAACTCTCCCGCAGCAGTGGCCCGCAGAGTACCAAGTGG 601
DB      3958 GGTCCCGGGTGTTCACGGATACTCTCTCCACAGTAGTGGCCCGCAGAGTTCAGGTGG 4017

QY      602 CCACCTGTCATGCTCCCGCAGCGGTAAAGACACCAAGGTCCCGCGCGCATAGCAG 661
DB      4018 CTCACCTTCATGCTCCCGCAGCGGGCCAAAGACCAAGGTCCCGCGCTGCATATGCA 4077

QY      662 CTCAGGGCTACAAGGTGCTGTCTCAACCCCTCGTGTGCTGCAACAATGGGGTGTGGTG 721
DB      4078 CTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTGTGCTGCAACAATGGGGTGTGGTG 4137

QY      722 CTTACATGTCCAAGGCCCATGGGATTGATCTTAACATCAGGACTGGGTGAGGACAATTA 781
DB      4138 CTTACATGTCCAAGGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA 4197

QY      782 CTACTGGCAGCCGATCACGTATTTCCACCTTACGGCAAGTTCCTTGGCAGCGGGGTGTT 841
DB      4198 CCACTGGCAGCCCATCACGTACTCCACTACAGGACCGGGGTGAGAACAAATTA 4257

QY      842 CAGGGGGTCTTATGACATAAATTTGTGACAGTGCCACTCCACGATGTCAACATCCA 901
DB      4258 CGGGGGCGCTTATGACATAAATTTGTGACAGTGCCACTCCACGATGTCAACATCCA 4317

QY      902 TCTTGGGCATTGGCACTGTCTTGACCAAGCAGAGACCGGGGGCGAGACTGACTGTGC 961
DB      4318 TCTTGGGCATCGGCACTGTCTTTGACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTGC 4377
```

QY 962 TCGCCACCGCTACCCCTCGGGCTCCGTCACGTGTCGCCCATCTTAACATCGAGGAGTTG 1021
Db 4378 TCGCCACCGCCACCCCTCGGGCTCCGTCACGTGTCGCCCATCTTAACATCGAGGAGTTG 4437
QY 1022 CTCTGTCCATACCGGAGAGATCCCTTTTATGGCAAGCTATTCCCTTGAAGCAATTA 1081
Db 4438 CTCTGTCCATACCGGAGAGATCCCTTTTATGGCAAGCTATTCCCTTGAAGCAATTA 4497
QY 1082 AGGGGGGAGACATCTCATCTCTCCCACTCAAGAAGAAGTGGACGAGCTCGCGCAA 1141
Db 4498 AGGGGGGAGACATCTCATCTCTCCCACTCAAGAAGAAGTGGACGAGCTCGCGCAA 4557
QY 1142 AACTGTGCGCTTGGCGCTCAATCCGCTGCTTACTACGGGGCTTGTATGTCTCCGTC 1201
Db 4558 AGCTGTGCGCATTTGGGCATCAATCCGCTGCTTACTACGGGGCTTGTATGTCTCCGTC 4617
QY 1202 TCCCGACCAAGTGTGCGCTGCTGCGCAACTGACGCGCTCATGACCGCTTTTACCG 1261
Db 4618 TCCCGACCAAGTGTGCGCTGCTGCGCAACTGACGCGCTCATGACCGCTTTTACCG 4677
QY 1262 CGGACTTCGATTCGGTGATAGACTGCAACACGCTGTGTCAACCCAGACAGTTCAGCC 1321
Db 4678 CGGACTTCGATTCGGTGATAGACTGCAACACGCTGTGTCAACCCAGACAGTTCAGCC 4737
QY 1322 TTGACCTTACCTTACCATTTAGACAAATCAAGCTTCCCGAGAGTGTCTCCCGTACTC 1381
Db 4738 TTGACCTTACCTTACCATTTAGACAAATCAAGCTTCCCGAGAGTGTCTCCCGTACTC 4797
QY 1382 AACGTCGGGTGAGACTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1441
Db 4798 AACGTCGGGTGAGACTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4857
QY 1442 AGCGTCTTCTGGCATTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
Db 4858 AGCGTCTTCTGGCATTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4917
QY 1502 CTTGTGTAGAGTTACGCGCGGAGACCAAGTTAGGCTACGAGTATACATGACACCC 1561
Db 4918 CTTGTGTAGAGTTACGCGCGGAGACCAAGTTAGGCTACGAGTATACATGACACCC 4977
QY 1562 CGGAGCTTCCGCTGTCAGACCATCTTGAATTTGGAGGGCTCTTTACGGGTCTCA 1621
Db 4978 CGGAGCTTCCGCTGTCAGACCATCTTGAATTTGGAGGGCTCTTTACGGGTCTCA 5037
QY 1622 CCCATAGACGCCACTTCTATCCAGACAAAGACAGTGGGAAACCTTCCCTATC 1681
Db 5038 CTATATAGATGCCACTTCTATCCAGACAAAGACAGTGGGAGAACCTTCCCTATC 5097
QY 1682 TGGTAGCGTACCAAGCCACCGTGTGCGTACAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
Db 5098 TGGTAGCGTACCAAGCCACCGTGTGCGTACAGCTCAAGCCCTCCCGCTCGTGGGACC 5157
QY 1742 AGATGTGAAGTGTGATCGTCTCAAGCCACCCCTCCATGGGCGCACCTCTGCTAT 1801
Db 5158 AGATGTGAAGTGTGATCGTCTCAAGCCACCCCTCCATGGGCGCACCTCTGCTAT 5217
QY 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTTGACGACCCAGTACCAAGTATATCA 1861
Db 5218 ACAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTTGACGACCCAGTACCAAGTATATCA 5277
QY 1862 TGACATGATGTGCGCTGACTGGAGGTGCTGACAGTACTGCGGTGCTGCTGGCGCG 1921
Db 5278 TGACATGATGTGCGCTGACTGGAGGTGCTGACAGTACTGCGGTGCTGCTGGCGCG 5337
QY 1922 TTCTGGCTGCTTTGGCGCGTATGCTCTATCCAGGCTGCTGCTGCTATAGTAGTAGGA 1981
Db 5338 TTCTGGCTGCTTTGGCGCGTATGCTCTATCCAGGCTGCTGCTGCTATAGTAGTAGGA 5397
QY 1982 TTGTCTTGTTCGGAAGCGCGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTTCG 2041
Db 5398 TCGTCTTGTTCGGAAGCGCGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTTCG 5457
QY 2042 ATGAATGGAAGAGTGCT 2059

Db 5458 ATGAGATGGAAGAGTGCT 5475

RESULT 6

US-08-440-549-9
; Sequence 9, Application US/08440549
; Patent No. 6312889
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,549
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9401 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 342..9374
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 366
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Arg."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 372
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 867
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1341
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature

/ LOCATION: 2148 /note= "This amino acid position
/ OTHER INFORMATION: can also be Ile."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 2883
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Asn."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3681
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3690
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4167
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Leu."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4323
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4701
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Tyr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4752
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 5970
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6183
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be His."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6186
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Cys."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6402
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7386
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7494
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7497
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ala."

/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7845
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 8409
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9102
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9327
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Pro."
/ US-08-440-549-9
Query Match 86.7%; Score 1786; DB 3; Length 9401;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 2 TGGCGCCTATCAGGCGCTATGCCAGCAGACAGAGGGGCTTTTGGGATGCATAATCACCA 61
DB 3418 TGGCGCCCATCAGCGGTACGCCACGACAGACAGAGGGGCTTCTAGGGTGCATAATCACCA 3477
QY 62 GCTTGACCGCGCGGACAAAACACAGGTGGAGGGTTCAGATTTCAGATCGTGTCAACTGCTG 121
DB 3478 GCCTAACTGCGCGGACAAAACCAAGTGGAGGGTGCAGATTGTGTCAACTGCTG 3537
QY 122 CCAGACTTTCTTGGCAACTGCTGCAATTAAACGGGGTGTGTGGAGTGTCTTACCATGGAGCCG 181
DB 3538 CCCAAACCTTCTCTGGCAAGTGCATCAATGGGGTGTGTGGAGTGTCTTACCATGGGCGG 3597
QY 182 GAACAGGACCATTTGGTGCCTGACCTAAGGGTCTCTTATCCAGATGTACACCAATGTGACC 241
DB 3598 GAACAGGACCATTCGGCTCACCACAGGGTCTGTCTATCCAGATGTATACCAATGTAGACC 3657
QY 242 AAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCGCGCTCATTAACACCATGCACTTGGC 301
DB 3658 AAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCGGCTCATTTGACACCCCTGCACTTGG 3717
QY 302 GCTCCTCGGACCTTTACCTGGTGCACAGGACGCGCATGTCTTCTGTGCGCGGAGCGG 361
DB 3718 GCTCCTCGGACCTTTACCTGGTGCACAGGACGCGCATGTCTTCTGTGCGCGGAGCGG 3777
QY 362 GTGATGCGAGGGGCGAGCTGCTTTGCGCCCGGCTATCTTACTTTGAAAGGCTCCTCGG 421
DB 3778 GTGATGCGAGGGGCGAGCTGCTGCTGCGCCCGGCTATTTCTTACTTTGAAAGGCTCCTCGG 3837
QY 422 GAGGCGCTCTGTGTGCTCCCGGAGGACATGCGCTAGGCATATTTCAGAGCGCGGATGCA 481
DB 3838 GGGGTGCGCTGTTGTGCGCCCGGCGGACGCGGTGGGCATATTTAGGCGCGGCTGTGCA 3897
QY 482 CCGGTGGAGTGGCTAAGGGCGGTGGAATTTCATCCCGGTAGAGAGCTTAGAGACAAACATGA 541
DB 3898 CCGGTGGAGTGGCTAAGGGCGGTGGAATTTCATCCCTGTGGAGAACCTTAGAGACAAACATGA 3957
QY 542 GGTCCCGGCTGTTCTCAGACAACTCTCCCGGACAGTGGCCCGGAGAGCTACCAAGTGG 601
DB 3958 GGTCCCGGCTGTTCTCAGAGTAATCTCTCTCCACAGTAGTGGCCCGGAGAGCTTCAGGTGG 4017
QY 602 CCACCTGCATGCTCCCGGAGCGGTAAAGAGCAAGGTCCCGGCGCATACGAG 661
DB 4018 CTCACCTCCATGCTCCCGGAGCGGCGCAAAAGCAAGGTCCCGGCTGCATATGAG 4077
QY 662 CTCAGGGCTAACAGGTGCTGGTGTCTCAACCCCTCGGTGTGCTGCAACATGGGCTTTGGTG 721
DB 4078 CTCAGGGCTAACAGGTGCTAGTACTCAACCCCTCTGTGTGCTGCAACATGGGCTTTGGTG 4137

QY 722 CTTACATGTCAGAGCCCATGGGATTGATCTTAACATCAGACTCGGGGTGAGGACAATTA 781
DB 4138 CTTACATGTCAGAGCCCATGGGATTGATCTTAACATCAGACTCGGGGTGAGGACAATTA 4197
QY 782 CTACTGGCAGCCCGATCAGCGTATTCCTACGCAAGTTCCTTGGCCGAGCGGGGTGTT 841
DB 4198 CCACTGGCAGCCCGATCAGCGTATTCCTACGCAAGTTCCTTGGCCGAGCGGGGTGCT 4257
QY 842 CAGGGGTGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGAGATGCAACATCCA 901
DB 4258 CGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGAGATGCAACATCCA 4317
QY 902 TCTTGGGATTTGGCACTGTCTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 961
DB 4318 TCTTGGGATTTGGCACTGTCTTGACCAAGCAGAGACTGCGGGGGGAGACTGCTGTTGTC 4377
QY 962 TCGCCACCGCTACCCCTCCGGGCTCCGTCACGTGCCCCATCCTAAACATCGAGAGGTTG 1021
DB 4378 TCGCCACCGGCCCCCTCCGGGCTCCGTCACGTGCCCCATCCTAAACATCGAGAGGTTG 4437
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAACAATTA 1081
DB 4438 CTCTGTCCACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAAGTAATCA 4497
QY 1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAAAGTGCAGAGCTCGCCGCAA 1141
DB 4498 AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAAAGTGCAGAGCTCGCCGCAA 4557
QY 1142 AACTGTTGCGGTTGGGCGTCAATCCGCTGCTTACTACCGGGCTTGTGATGTGCTGCA 1201
DB 4558 AGTGGTGCATTTGGGATCAATCCGCTGCTTACTACCGGGCTTGTGATGTGCTGCA 4617
QY 1202 TCCGACAGTGGTGGTGTGCTGCTGCTGGCACTGACGCTCATGACCGGCTTTTACCG 1261
DB 4618 TCCGACAGCGGGATGTTGCTGCTGCTGGCAACCGATGCCCTCATGACCGGCTTACCG 4677
QY 1262 GCGACTTCGATTGGTGATAGACTGCAACAGTGTGTCAACAGACAGTGCAGTTACGCC 1321
DB 4678 GCGACTTCGATTGGTGATAGACTGCAACAGTGTGTCAACAGACAGTGCAGTTACGCC 4737
QY 1322 TTGACCTTACCTTACCATTTGAGACATCAGCTTCCGAGGATGCTGTCCCGTACTC 1381
DB 4738 TTGACCTTACCTTACCATTTGAGACATCAGCTTCCCGAGGATGCTGTCCCGCACTC 4797
QY 1382 AACGTGGGGTAGGACTGGCAGAGGGAAGCAGGATCTACAGATTTGTGCAACGGGG 1441
DB 4798 AACGTGGGGTAGGACTGGCAGAGGGAAGCAGGATCTACAGATTTGTGCAACGGGG 4857
QY 1442 AGCGTCTCTTGGCATGTTTGAATGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
DB 4858 AGCGCTCTCTGCGCATGTTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4917
QY 1502 CTTGGTATGAGTTTACCGCCCGCAGACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1561
DB 4918 CTTGGTATGAGTTTACCGCCCGCAGACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 4977
QY 1562 CGGACTTCCGTTGTCAGAACCATCTGAAATTTTGGGAGGCGCTTTTACGGGCTCA 1621
DB 4978 CGGAGCTTCCGTTGTCAGAACCATCTGAAATTTTGGGAGGCGCTTTTACGGGCTCA 5037
QY 1622 CCACATAGAGCCCACTTCTTATCCAGACCAAGCAGAGTGGGGAAGCTTCCCTTATC 1681
DB 5038 CTATATAGATGCCACTTCTTATCCAGACCAAGCAGAGTGGGGAAGCTTCCCTTATC 5097
QY 1682 TGGTAGGATACCAAGCCACCGTGTGCTGAGCTCAAGCCCTTCCCGCTGCTGGGACC 1741
DB 5098 TGGTAGGATACCAAGCCACCGTGTGCTGAGCTCAAGCCCTTCCCGCTGCTGGGACC 5157
QY 1742 AGATGTGGAAGTGTGATCGCTTCAAGCCCAAGCTCCATGGGCAACACCTTGTAT 1801
DB 5158 AGATGTGGAAGTGTGATCGCTTCAAGCCCAAGCTCCATGGGCAACACCTTGTAT 5217

QY 1802 ATAGACTGGGGCTGTCCAGAAATGAAGTCACTCACTGACGCAACCCAGTCAACCAAGTATATCA 1861
DB 5218 ACAGACTGGGGCTGTTCAGAAATGAAGTCACTCACTGACGCAACCCAGTCAACCAAGTATATCA 5277
QY 1862 TGACATCTATGTGGCGTACCTGGAGTCTGTCAGAGTACCTGGTGTCTGTTGGCGGG 1921
DB 5278 TGACATCTATGTGGCGTACCTGGAGTCTGTCAGAGTACCTGGTGTCTGTTGGCGGG 5337
QY 1922 TTCTGGCTGTCTTGGCGCGCTATTGCTCTATCCACAGGCTGGTGTCTATAGTAGTAGGA 1981
DB 5338 TCCTGGCTGTCTTGGCGCGCTATTGCTCTATCCACAGGCTGGTGTCTATAGTAGTAGGA 5397
QY 1982 TTGTCTTGTCCGAAAGCCGCAATCATACCGGACAGGAGTCTCTTACCGGAGTTTCG 2041
DB 5398 TCCTTGTCTCCGGAAGCCGCAATCATACCGGAGGAGTCTCTTACCGAGAGTTTCG 5457
QY 2042 ATGAATGGAAGAGTGCT 2059
DB 5458 ATGAGATGGAAGAGTGCT 5475

RESULT 7

US-08-823-895A-25

; Sequence 25, Application US/08823895A

; Patent No. 6433159

; GENERAL INFORMATION:

; APPLICANT: Kevin P. Anderson

; TITLE OF INVENTION: Compositions And Methods For

; TITLE OF INVENTION: Treatment Of Hepatitis C Virus-Associated Diseases

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Jane Massey Licata, Esq.

; STREET: 66 E. Main Street

; CITY: Marlton

; STATE: NJ

; COUNTRY: USA

; ZIP: 08053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE

; COMPUTER: IBM 486

; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/823,895A

; FILING DATE: March 17, 1997

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/453,085

; FILING DATE: May 30, 1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/945,289

; FILING DATE: September 10, 1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Jane Massey Licata

; REGISTRATION NUMBER: 32,257

; REFERENCE/DOCKET NUMBER: ISPH-0203

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (609) 779-2400

; TELEFAX: (609) 810-1454

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9401

; TYPE: Nucleic

; STRANDEDNESS: Single

; TOPOLOGY: Linear

; ANTI-SENSE: NO

; US-08-823-895A-25

Query Match 86.7%; Score 1786; DB 3; Length 9401;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1889; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTATCACGGCTATGCCCGCTATGCCAGCAGCAAGGGGCGCTTTTGGGATGCATATACCA 61

[illegible]

4498	AGGGGGGAGACATCTCATCTCTCTGTCTATTCAAAGAAAGTGCAGCAACTCGCCGCA	4555
1142	AATGGTTCGGTTGGCGTCAATGCGGTGGCTTACTACCGCGGCTTGATGTGTCGTCA	1201
4558	AGCTGGTTCGATTTGGGCATCAATGCGGTGGCTTACTACCGCGTCTTGACGTTCGGTCA	4617
1202	TCCGACCAAGTGTGAGCTTGTTGTCGTGGCAACTACGCCCTCATGACCGGCTTTACCG	1261
4618	TCCCGACCAAGCGCGATGTTGTCGTGGCAACCGATGCCCTCATGACCGGCTATACCG	4677
1262	GGACTTCGATTCCGTTAGACTGCAACACGTTGTCAACCCAGACAGTCGACTTCAGCC	1321
4678	GGACTTCGACTCGGTATAGACTGCAATACGTTGTCAACCCAGACAGTCGATTCAGCC	4737
1322	TTGACCTACCTTTCACATTTAGACAATCACGTTTCCCAAGGATGTGTCTCCGTACTC	1381
4738	TTGACCTACCTTTCACATTTAGACAATCACGTTTCCCAAGGATGTGTCTCCCGCACTC	4797
1382	AACGTGGGGTAGGACTGGCAGAGGAGCCAGGCATCTACAGATTTGTGGCACCCGGGG	1441
4798	AACGTGGGGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCCGGGG	4857
1442	AGCGTCTCTGCGCATGTTTGAATCGTCTGTCTCTCGAGTGTCTATGACCGGGTTGTG	1501
4858	AGCGCCCTTCGGCATGTTTGCATCTCGTCCGTCTCTGTGAGTGTATGACGAGGCTGTG	4917
1502	CTTGGTATGACTTACGCCCGCGAGACCAAGTTAGGCTACGAGCATCATGAACACCC	1561
4918	CTTGGTATGACTTACGCCCGCGAGACTACAGTTAGGCTTACGAGGCTCATGAACACCC	4977
1562	CGGGACTTCCCGGTGCCAAGACATCTTGAAATTTTGGGAGGGCGTCTTTTACGGGTCTCA	1621
4978	CGGGCTTCCCGGTGCCAGGACATCTTGAAATTTTGGGAGGGCGTCTTTTACGGGCTCA	5037
1622	CCACATAGACGCCACTTCCCTATCCAGACAAGACAGAGTGGGGGAAACCTTTCCCTATC	1681
5038	CTCATATAGATGCCACTTTCTATCCAGACAAGCAGAGTGGGGAGAACCTTTCTTACC	5097
1682	TGTFAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCTCCCGTCTGGGACC	1741
5098	TGTFAGCGTACCAAGCCACCGTGTGCGCTAGGCGTCAAGCCCTCTCCCGTCTGGGACC	5157
1742	AGATGTGGAAGTGTGATCCGTTCTCAAGCCCAACCTTCCATGGGCCAAACCTTCTGCTAT	1801
5158	AGATGTGGAAGTGTGATTCGCTCAAGCCCAACCTTCCATGGGCCAAACCTTCTGCTAT	5217
1802	ATAGACTGGGCGGTGTCCAGAAATGAAGTCAACCTTGACGCAACCAAGTCAATATCA	1861
5218	ACAGACTGGGCGGTGTTTCAATGAATCAACCTTGACGCAACCCAGTCAACAAATACATCA	5277
1862	TGACATGTATGTGCGCTGAACCTGAGAGTCTGTACAGAGTACTGGTGTCTCGTTGGCGCG	1921
5278	TGACATGCATGTGCGCGGACTGGAGGTCTGTCAGAGCACCTGGGTGTCTCGTTGGCGCG	5337
1922	TTCTGGCTGTCTGGCGCGCTATTGCTATTCACAGGCTCGTGGTCTATAGTAGTAGGA	1981
5338	TCCTGGCTGTCTTTGGCGCGGTATTGCTGTCAACAGGCTCGTGGTCTATAGTAGTAGGA	5397
1982	TTGTCTTGTCCGGAACCGCGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTCG	2041
5398	TCGTCTTGTCCGGAACCGCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTCG	5457
2042	ATGAAATGGAAGAGTGCT	2059
5458	ATGAGATGGAAGAGTGCT	5475

```

RESULT 8
PCT-US91-02225-9
; Sequence 9, Application PC/TUS9102225
; GENERAL INFORMATION:
; APPLICANT: HOUGHTON, MICHAEL
; APPLICANT: CHOO, QUI-LIM

```

APPLICANT: KUO, GEORGE
TITLE OF INVENTION: COMBINATIONS OF HEPATITIS C VIRUS
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 545 Middlefield Road, Suite 200
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02225
FILING DATE: 19910329
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CIOTTI, THOMAS E.
REGISTRATION NUMBER: 21,013
REFERENCE/DOCKET NUMBER: 2300-0101.44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 327-7250
TELEFAX: (415) 327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 9401 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
PCT-US91-02225-9

Query Match 86.6%; Score 1785.2; DB 5; Length 9401;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 1883; Conservative 6; Mismatches 169; Indels 0; Gaps 0;

Qy	2	TGGCGCTATCAGCGCTATGCCAGAGCAAGGGGCTTTGGGATGCAATACCA	61
Db	3418	TGGCGCCCATCAGCGCTATGCCAGAGCAAGGGGCTTTGGGATGCAATACCA	3477
Qy	62	GCTTGACGGCGGGGCAAAACCAGGTGGAGGTGAGTTTCAGATGTCCTCACTGCTG	121
Db	3478	GCCTAACTGGCGGGGCAAAACCAGGTGGAGGTGAGTTTCAGATGTCCTCACTGCTG	3537
Qy	122	CCAGACTTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTTACCATGGAGCG	181
Db	3538	CCCAAACTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTTACCAGGGGCG	3597
Qy	182	GAACAAGGACATTTGGCTACCTAAGGCTCTGTATCCAGATGTACCAATGTGAAC	241
Db	3598	GAACAGGACCATCGCGTACCCCAAGGCTCTGTATCCAGATGTATACCAATGTAGAC	3657
Qy	242	AAGACCTGTAGGTGGCGCGCTCCCAAGTGCCCGCTCATTAACCACTGCACTTGG	301
Db	3658	AAGACCTGTGGGTGGCGCGCTTCGCAAGGTATCCCGCTCATTAACCACTTGG	3717
Qy	302	GCTCTCGGACTTTTACCTGTGTACAGGACGCGCGATGTCTTCTGTGGCGCGAGGG	361
Db	3718	GCTCTCGGACTTTTACCTGTGTACAGGACGCGCGATGTCTTCTGTGGCGCGAGGG	3777
Qy	362	GTGATGGCAGGGGAGCGCTGTCTTGGCGCGCGCTATCTTACTTGAAGGCTCTCGG	421
Db	3778	GTGATGACAGGGGAGCGCTGTCTTGGCGCGCGCTATCTTACTTGAAGGCTCTCGG	3837
Qy	422	GAGGCCCTCTGCTGTGCGCGCGAGACATGCGGTAGGCATATTCAGAGCGCGGTATGCA	481
Db	3838	GGGGTCCGCTGTGTGTGCGCGCGAGACATGCGGTAGGCATATTCAGAGCGCGGTATGCA	3897

Qy	482	CCCGTGGAGTGGCTAAGCGCGGTGACATTTCATCCCGGTAGAGAGCTTAGAGCAACCAATGA	541
Db	3898	CCCGTGGAGTGGCTAAGCGCGGTGACATTTCATCCCGGTAGAGAACTTAGAGCAACCAATGA	3957
Qy	542	GGTCCCGGTGTCTCAGACAACTCTCCACACAGCAGTGGCCCGAGAGCTTACCAAGTGG	601
Db	3958	GGTCCCGGTGTCTCAGAGTAACCTCTCTCACAGTAGTGGCCCGAGAGCTTCCAGGTGG	4017
Qy	602	CCACCTGTCATGCTCCACCGGCAAGCGGTAAAGACCAACAAAGTCCCGGCCCATACGAG	661
Db	4018	CTACCTCCATGCTCCACAGGAGCGGCAAAAGCAACAAAGTCCCGGCTGCATATGAG	4077
Qy	662	CTCAGGCTACAAGGTGCTGGTCTCAACCCCTCCGTGTGTGCAACAAATGGCTTTGGTG	721
Db	4078	CTCAGGCTACAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACAACTGGCTTTGGTG	4137
Qy	722	CTTACATGTCNAAGGCTCATGGGATCGATCTTAACATCAGAGCTGGGTGAGCAATTA	781
Db	4138	CTTACATGTCNAAGGCTCATGGGATCGATCTTAACATCAGAGCTGGGTGAGCAATTA	4197
Qy	782	CTACTGGCAGCCGATCACGTTATTCACCTACGGCAAGTTCTCTTCCGACCGCGGTGTT	841
Db	4198	CCACTGGCAGCCGATCACGTTATTCACCTACGGCAAGTTCTCTTCCGACCGCGGTGCT	4257
Qy	842	CAGGGGTGCTTATGACATAAATATTTGTGACGAGTCCACTCCACGAGTCAACATCCA	901
Db	4258	CGGGGGCGCTTATGACATAAATATTTGTGACGAGTCCACTCCACGAGTCCACATCCA	4317
Qy	902	TCCTGGCAATGGCACTGCTCTTGACCAAGCAGAGACCGGGGGCGAGACTGACTGTGC	961
Db	4318	TCCTGGCAATGGCACTGCTCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGCTGTGC	4377
Qy	962	TCGCCACCGCTACCCCTCCCGGCTCCGTCACCTGTCGCCCTCTTAAATCAGAGAGTTG	1021
Db	4378	TCGCCACCGCAACCCCTCCCGGCTCCGTCACCTGTCGCCCTCTTAAATCAGAGAGTTG	4437
Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGSCAAGGCTATTTCCTCTTGAACAATTA	1081
Db	4438	CTCTGTCCACTACCGGAGAGATCCCTTTTATGSCAAGGCTATTTCCTCTTGAAGTAATTA	4497
Qy	1082	AGGGGGGAGACATCTCTTCTGCCACTCAAAAGAGAGTGGCAGAGCTCCCGCAA	1141
Db	4498	AGGGGGGAGACATCTCTTCTGCCACTCAAAAGAGAGTGGCAGAGCTCCCGCAA	4557
Qy	1142	AACGTGTCGGTGGGGTCAATGCGGTGCTTACTACCGCGCTTGTATGTGTCGGTCA	1201
Db	4558	AGCTGTGCGATTGGGGTCAATGCGGTGCTTACTACCGCGCTTGTATGTGTCGGTCA	4617
Qy	1202	TCCGACACAGTGGTGTGCTGCTGGCAACTGACCGCTCATGACCGGCTTTACCG	1261
Db	4618	TCCGACACAGCGCGATGTTGTGCTGCTGGCAACCGATGCCCTCATGACCGGCTTACCG	4677
Qy	1262	GGACTTCGATTCGGTATAGCTGCAACAGTGTGTGCCAGACAGTGCAGCTTCAGCC	1321
Db	4678	GGCACTTCGATTCGGTATAGCTGCAACAGTGTGTGCCAGACAGTGCAGCTTCAGCC	4737
Qy	1322	TTGACCCCTACCTTACCAATGAGCAATCAAGTTCGCCAGGATGCTGTCTCCCGTACTC	1381
Db	4738	TTGACCCCTACCTTACCAATGAGCAATCAAGTTCGCCAGGATGCTGTCTCCCGTACTC	4797
Qy	1382	AAAGTTCGGGTAGGACTGGCAGAGGGAAGCAGGCATCTACAGATTTGTGGCACCGGGG	1441
Db	4798	AAAGTTCGGGTAGGACTGGCAGAGGGAAGCAGGCATCTACAGATTTGTGGCACCGGGG	4857
Qy	1442	AGGCTCTCTTGGGATGCTGTGCTGTCTTCTCGAGTGTATGACGCGGGTGTG	1501
Db	4858	AGGCTCTCTTGGGATGCTGTGCTGTCTTCTCGAGTGTATGACGCGGGTGTG	4917
Qy	1502	CTTGGTATGAGCTTACCGCGCGAGACCAAGTTAGGCTACGAGCATATGATGAACACC	1561
Db	4918	CTTGGTATGAGCTTACCGCGCGAGACCAAGTTAGGCTACGAGCATATGATGAACACC	4977

QY 1562 CGGACTCCCGTGTGCCAAGACCATCTTGAATTTGGAGGGCGCTTTTACGGGTCTCA 1621
Db CGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGAGGGCGCTTTTACAGGCTCA 5037
QY 1622 CCACATAGAGCCCACTTCTATCCAGACAAAGAGAGAGTGGGAAACCTTCCCTATC 1681
Db CTCATATAGATGCCCACTTCTATCCAGACAAAGAGAGAGTGGGAGAACCTTCTTACC 5097
QY 1682 TGGTAGGTTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
Db TGGTAGGTTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 5157
QY 1742 AGATGTGGAAGTGTGATCCGCTCAAGCCCACTTCCATGGGCGCAACCACTCTGCTAT 1801
Db AGATGTGGAAGTGTGATCCGCTCAAGCCCACTTCCATGGGCGCAACCACTCTGCTAT 5217
QY 1802 ATAGACTGGGCGGTGTCCAGATGAAGTCACTTCCAGCGCACCCAGTCCCAAAATACATCA 1861
Db ACAGACTGGGCGGTGTCCAGATGAAGTCACTTCCAGCGCACCCAGTCCCAAAATACATCA 5277
QY 1862 TGACATGTATGTGGCTGACTTGAGGTCTGACGAGTACCTGGGTGCTGTTGGCGGG 1921
Db TGACATGTATGTGGCTGACTTGAGGTCTGACGAGTACCTGGGTGCTGTTGGCGGG 5337
QY 1922 TTCTGGCTGCTTTGGCGCGTATTGCCCTATCCAGGCTGCGTGTGCTAGTAGTAGGA 1981
Db TCCTGGCTGCTTTGGCGCGTATTGCCCTATCCAGGCTGCGTGTGCTAGTAGTAGGA 5397
QY 1982 TTGCTTGTCCGAAGCCGGCAATCATACCAGAGGAGTCTTCTACCGGAGTTGCG 2041
Db TCCTGTTGTCCGAAGCCGGCAATCATACCAGAGGAGTCTTCTACCGGAGTTGCG 5457
QY 2042 ATGAAATGGAAGAGTGCT 2059
Db ATGAGATGGAAGAGTGCT 5475

RESULT 9

US-08-444-818-65
; Sequence 65, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 6785 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..6785
; US-08-444-818-65

Query Match 86.6%; Score 1784.4; DB 3; Length 6785;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 2 TGGCGCTATACAGGGCTATGCCAGACAGACAAAGGGGCCCTTTGGATGCAATACCA 61
Db 1204 TGGCGCCATACAGGGCTATGCCAGACAGACAAAGGGGCCCTTCTAGGGTGCAATACCA 1263
QY 62 GCTTGACCGCGGAGACAAACCAAGGTGAGGGTGAGGTTGAGATCGTCAACTGCTG 121
Db 1264 GCTTAACTGGCGGAGACAAACCAAGGTGAGGGTGAGGTTGAGATCGTCAACTGCTG 1323
QY 122 CCAGACTTTCTTGGCAACCTGCAATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCG 181
Db 1324 CCAGAACCTTCTGGCAACCTGCAATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCG 1383
QY 182 GAACAGGACCAATGGGTCACTAAGGTCTCTGTTATCCAGATGTACCAATGGAGCC 241
Db 1384 GAACAGGACCAATGGGTCACTAAGGTCTCTGTTATCCAGATGTACCAATGGAGCC 1443
QY 242 AAGACCTGCTAGCGTGGCGCGCTCCCAAGGTGCCCGCTATTAAACCATGCACCTGGC 301
Db 1444 AAGACCTGCTAGCGTGGCGCGCTCCCAAGGTGCCCGCTATTAAACCATGCACCTGGC 1503
QY 302 GCTCTCGGACCTTTACCTGCTCAGAGGACCGCCGATGCTATTCCTGTGCCCGGAGCG 361
Db 1504 GCTCTCGGACCTTTACCTGCTCAGAGGACCGCCGATGCTATTCCTGTGCCCGGAGCG 1563
QY 362 GTGATGGAGGGGAGCGCTGCTTTGCCCGCGGCTATCTCTTACTTGAAGGCTCTCGG 421
Db 1564 GTGATGGAGGGGAGCGCTGCTTTGCCCGCGGCTATCTCTTACTTGAAGGCTCTCGG 1623
QY 422 GAGGGCTCTGCTGTGCCCCGAGACATGCGTAGGCAATATTCAGAGCCGCGGTATGCA 481
Db 1624 GAGGGCTCTGCTGTGCCCCGAGACATGCGTAGGCAATATTCAGAGCCGCGGTATGCA 1683
QY 482 CCCGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCGGTAGAGAGCTTAGAGACCAATGA 541
Db 1684 CCCGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCGGTAGAGAGCTTAGAGACCAATGA 1743
QY 542 GGTCCCGGCTTCTCAGACAACTCTCTCCACAGAGAGTGCCTCCAGAGCTACCAAGTG 601
Db 1744 GGTCCCGGCTTCTCAGACAACTCTCTCCACAGAGAGTGCCTCCAGAGCTACCAAGTG 1803
QY 602 CCCACTGCTATCTCCACCGGAGCGGTAAAGACCAAGGTGCCCGCGCATACGAG 661
Db 1804 CTCACCTCATGCTCCACAGGAGCGGCAAAAGCAACAGGTCCCGGCTGATATGAG 1863
QY 662 CTCAGGGCTTACAAGGTGCTGGTGTCAACCCCTCGCTGCTGCAACAATGGGCTTTGGTG 721
Db 1864 CTCAGGGCTTACAAGGTGCTGGTGTCAACCCCTCGCTGCTGCAACAATGGGCTTTGGTG 1923
QY 722 CTTACATGTCCAGGCCCATGGGATGATCTTAAATCAGAGCTGGGTGAGGACAATTA 781
Db 1924 CTTACATGTCCAGGCCCATGGGATGATCTTAAATCAGAGCTGGGTGAGGACAATTA 1983
QY 782 CTACTGGCAGCCCGATCAGTATTTCACCTACGGCAAGTTCCTTCCGACGCGGCTGTT 841
Db 1984 CCACTGGCAGCCCGATCAGTATTTCACCTACGGCAAGTTCCTTCCGACGCGGCTGCT 2043
QY 842 CAGGGGTGCTTATGACATAATAATTTGTGAGAGTGCCACTCCACGATGCAACATCA 901
Db 842 CAGGGGTGCTTATGACATAATAATTTGTGAGAGTGCCACTCCACGATGCAACATCA 901

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885

INFORMATION FOR SEQ ID NO: 137:
SEQUENCE CHARACTERISTICS:
LENGTH: 8987 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..8985
US-08-444-818-137

Query Match 86.6%; Score 1784.4; DB 3; Length 8987;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy	2	TGGCGCCTATCAGCGCCTATGCCAGCAGACAAAGGGCCCTTTGGATGCAATATCACCA	61
Db	3077	TGGCGCCCATCACGGCGTACGCCAGCAGACAAAGGGCCCTCTAGGGTGCATATCACCA	3136
Qy	62	GCTTGACCGGCGGGACAAACAGGTGGAGGTGAGTTTCAGATCGTGTCAACTGCTG	121
Db	3137	GCCTAACTGGCGGGACAAACAGGTGGAGGTGAGTTTCAGATCGTGTCAACTGCTG	3196
Qy	122	CCCAGACTTCTTGGCAACTGCAATTAACGGGGTGTGTGGACTGTCTACCATGGAGCCG	181
Db	3197	CCCAAACTTCTGGCAACTGCAATTAACGGGGTGTGTGGACTGTCTACCATGGAGCCG	3256
Qy	182	GAACAGACCACTTGGCTACCTAAGGTCCTGTATCCAGATGTACCAATGTGGACC	241
Db	3257	GAACAGAGACCACTGGCTACCTAAGGTCCTGTATCCAGATGTATACCAATGTAGACC	3316
Qy	242	AAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGGC	301
Db	3317	AAGACCTTGTGGCTGGCCCGCTCCGAGGTAGCGGCTCATTAACACCATGCACTTGGC	3376
Qy	302	GCTCCTCGGACCTTTACCTGTGTACAGAGCACGCGCGATGTCTTCTGTGCGCGACGGG	361
Db	3377	GCTCCTCGGACCTTTACCTGTGTACAGAGCACGCGCGATGTCTTCTGTGCGCGCGGG	3436
Qy	362	GTGATGGCAGGGGAGCGCTGCTTTGCGCCCGGCGCTATCTTACTTGAAGGCTCTCTGG	421
Db	3437	GTGATAGCAGGGGAGCGCTGCTTGTGCGCCCGGCGCTATCTTACTTGAAGGCTCTCTGG	3496
Qy	422	GAGCGCTCTGCTGTGCCCGCAGACATGCGGTAGGCATATTACAGAGCCGCGGTATGCA	481
Db	3497	GGGGTCCGCTGTTGTGCCCGGCGGACGCGGTGGGCATATTAGGCGCGCGGTGTGCA	3556
Qy	482	CCCGTGGAGTGGCTAAGCGGTGGACTTTTATCCCGGTAGAGAGCTTAGAGACAAACCATGA	541
Db	3557	CCCGTGGAGTGGCTAAGCGGTGGACTTTTATCCCGGTAGAGAGCTTAGAGACAAACCATGA	3616
Qy	542	GGTCCCGGTGTTCTCAGACAACTCTCCCGCAGCAGAGTCCCGCAGAGCTTACCAAGTGG	601
Db	3617	GGTCCCGGTGTTCTCAGACAACTCTCTCCCGCAGAGTGTCCCGCAGAGCTTCCAGGTGG	3676

Qy	602	CCACCTTGATGCTCCACCGGCGAGCGGTAAAGACACCAAGGTCCCGCGCATACGCGAG	661
Db	3677	CTCACCTCCATGCTCCACAGGCGAGCGGCAAAACCAAGGTCCCGGCTCATATGCGAG	3736
Qy	662	CTCAGGGCTACAGGTGCTGCTCAACCCCTCGTGTGCTGCAACAATGGGCTTTGGTG	721
Db	3737	CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGTGCAACACTGGGCTTTGGTG	3796
Qy	722	CTTACATGTCCAAGGCCATGGGATTGATCTTAACATCAGGACTGGGTGAGGACAATTA	781
Db	3797	CTTACATGTCCAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGACAATTA	3856
Qy	782	CTACTGGCAGCCCGATCACGTATTCCACCTACGGCAAGTTCTCTTGGCAGCGCGGTGTT	841
Db	3857	CACTGGCAGCCCGATCACGTACTCCACCTACGGCAAGTTCTCTTGGCAGCGCGGTGCT	3916
Qy	842	CAGGGGTGCTTATGACATTAATTTGTGACGATGCCACTCCACGGATCAACATCCA	901
Db	3917	CGGGGGCGCTTATGACATTAATTTGTGACGATGCCACTCCACGGATCAACATCCA	3976
Qy	902	TCCTGGGCATTGGCAGCTGCTCTTGACCAAGCAGAGACCGGGGGCGAGACTGACTGTC	961
Db	3977	TCCTGGGCATCGGCAGCTGCTCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTC	4036
Qy	962	TGCGCACCGCTTACCCCTCCGGGCTCCGTCACTGTGCCCCCATCTTAACATCAGGAGGTTG	1021
Db	4037	TGCGCACCGGCACCCCTCCGGGCTCCGTCACTGTGCCCCCATCTTAACATCAGGAGGTTG	4096
Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTTCCCTTTGAAGCAATTA	1081
Db	4097	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTTCCCTTTGAAGTAATCA	4156
Qy	1082	AGGGGGGAGACATCTCATCTTGTGCCACTCAAGAGAGAGTGCAGAGCTCCCGCAA	1141
Db	4157	AGGGGGGAGACATCTCATCTTGTTCATTAAGAGAGAGTGCAGAGCTCCCGCAA	4216
Qy	1142	AACTGGTGGGCTCAATGCCGCTTACTACCGCGCTTTGATGTGCTCCGTCA	1201
Db	4217	AGCTGGTGGGCTCAATGCCGCTTACTACCGCGCTTTGATGTGCTCCGTCA	4276
Qy	1202	TCCGACAGTGGTGAAGTGTGCTGTGGCAACTGACGCGCTCATGACCGGCTTTACCG	1261
Db	4277	TCCGACAGCGGCGATGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGCTATACCG	4336
Qy	1262	GCGACTTCGATTGGGTATAGACTGCAACAGTGTGTCAACAGACAGTGCAGCTTCAGCC	1321
Db	4337	GCGACTTCGACTGGGTATAGACTGCAATACGTGTGTCAACAGACAGTGCAGCTTCAGCC	4396
Qy	1322	TTGACCTTACCTTCAACATTTAGACAAATCACGCTTCCCGAGGATGCTGTCTCCGCTACTC	1381
Db	4397	TTGACCTTACCTTCAACATTTAGACAAATCACGCTTCCCGAGGATGCTGTCTCCGCTACTC	4456
Qy	1382	AAGTTCGGGTAGGACTGGCAGAGGAGGAGCGAGCATCTAAGATTTGTGCGACCGGGGG	1441
Db	4457	AAGTTCGGGTAGGACTGGCAGAGGAGGAGCGAGCATCAACAGATTTGTGCGACCGGGGG	4516
Qy	1442	AGCGTCTTCTGGCATGTTTGAATCGTCTCTCTCGGAGTGTATGACGCGGGTGTG	1501
Db	4517	AGCGTCTTCTGGCATGTTTGAATCGTCTCTCTCGGAGTGTATGACGCGGGTGTG	4576
Qy	1502	CTTGGTATGAGCTTACCGCCCGCAGACCAACAGTTAGGCTACGAGCATATACATGAACACCC	1561
Db	4577	CTTGGTATGAGCTACCGCCCGCAGACCAACAGTTAGGCTACGAGCATATACATGAACACCC	4636
Qy	1562	CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACGGGTCTCA	1621
Db	4637	CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCTCTCA	4696
Qy	1622	CCACATAGCGCCACTTCTCTATCCAGACAAAGCAGAGTGGGGAAACCTTCCCTATC	1681
Db	4697	CTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGAAACCTTCCCTATC	4756

QY 1682 TGCTAGCTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGACC 1741
Db |||||
QY 1742 AGATGGAAGTCTTGTATCCGCTCAAGCCCACTCCATGCGGCCCAACAACCTCTCTAT 1801
Db |||||
QY 1802 ATAGACTGGCGCTGTCCAGATGAAGTCACTCCCTGAGCGCACCCAGTCAACCAATATCA 1861
Db |||||
QY 1862 TGACATGTATGTCCGCTGACCTGGAGTCTGACAGTACCTGGGTGCTGTTGGCGGCG 1921
Db |||||
QY 1922 TTCTGCTGCTTTGGCGCGGTATGCTATCCAGGCTGCTGCTCATAGTAGTAGGA 1981
Db |||||
QY 1982 TTCTGCTGCTGCGAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTG 2041
Db |||||
QY 2042 ATGAAATGGAAGAGTCT 2059
Db |||||
QY 5117 ATGAGATGGAAGAGTCT 5134
Db |||||

RESULT 12

US-08-444-818-122
; Sequence 122, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-444-818-122

Query Match 86.6%; Score 1784.4; DB 3; Length 9185;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
QY 2 TGGCGCTATCAAGCGCTATGCCAGCAGACAAAGGGGCCCTTTGGGATGCAATATCA 61
Db TGGCGCCCATACGCGCTACGCCAGCAGACAAAGGGGCCCTCTAGGGTGCAATATCA 3455
QY 62 GCTTGACCGCGCGGACAAACACAGGTGAGGGTGGAGTTCAGATCGGTGTAACCTGCTG 121
Db GCTTAACCTGCGCGGACAAACACAGGTGAGGGTGGAGTTCAGATCGGTGTAACCTGCTG 3515
QY 122 CCCAGACTTTCTTGGCAACCTGCTATTAACGGGGTGTGTTGGAGTGTCTACCAATGGAGCCG 181
Db CCCAAACCTTCTGCGCAACGCTGATCAATGAGGGTGTGCTGAGCTGTCTACCAAGGGGCG 3575
QY 182 GAACAGGACCAATGCGGTCACTAAGGTTCCTGTTATCCAGATGTACACCAATGTGAGCC 241
Db GAACAGGACCAATGCGGTCACTAAGGTTCCTGTTATCCAGATGTACACCAATGTGAGCC 3635
QY 242 AAGACCTCTAGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCAATGCACCTTGGC 301
Db AAGACCTCTGGCTGGCGCGCTCCCAAGGTAGCGCTCATTTGACACCTGCACCTTGGC 3695
QY 302 GCTCTCGGACCTTTACCTGGTCAAGGCAAGCGGATGCTATCTGTGCGCGCAGCGG 361
Db GCTCTCGGACCTTTACCTGGTCAAGGCAAGCGGATGCTATCTGTGCGCGCGCGG 3755
QY 362 GTGATGGAGGGGACCGCTGCTTTCGCGCGGCTATCTTACTTTGAAGGCTCTCGG 421
Db GTGATAGCAGGGGACCGCTGCTTTCGCGCGGCTATCTTACTTTGAAGGCTCTCTCGG 3915
QY 422 GAGGCGCTCTGCTGTGCGCGCAGGACATGCGGTAGGCATATTCAGAGCGCGGCTATGCA 481
Db GGGGTCGCTGTGTCGCGCGGCGCAGCGGTGGCATATTTAGGGCGCGGCTGCA 3875
QY 482 CCGGTGAGTGGCTAAGCGGTGGATTTATCCCTGTGAGAGCTTTAGAGCAACCATGA 541
Db CCGGTGAGTGGCTAAGCGGTGGATTTATCCCTGTGAGAGCTTTAGAGCAACCATGA 3935
QY 542 GGTCCCGGGTGTCTCAGACAACTCTCCCAAGCAGTGGCCAGAGCTACCAAGTGG 601
Db GGTCCCGGGTGTCTCAGACAACTCTCTCCCAAGTGGCCAGAGCTTTCCAGGTGG 3995
QY 602 CCACCTGCTATCTCCCAAGCGGTAAAGACCAAGGTCCCGCGCGCATACGCGAG 661
Db CTCACTCCATGCTCCCAAGCGGTAAAGACCAAGGTCCCGCGCGCATATGCGAG 4055
QY 662 CTCAGGGCTACAGGTGTGCTGCTCAACCCCTCGTGTGCTGCAACAATGGGGTTGGTG 721
Db CTCAGGGCTATAAGGTGTGCTGCTCAACCCCTCTGTGCTGCAACAATGGGGTTGGTG 4115
QY 722 CTTACATGTCCAGGCGCATGGATTTGATCCTAACATCAGGAGTCCGGGTGAGCAATTA 781
Db CTTACATGTCCAGGCGCATGGATTTGATCCTAACATCAGGAGTCCGGGTGAGCAATTA 4175
QY 782 CTACTGGCAGCCCGATCAGCTATTCACCTTACGCAAGTTCCTTTCGCGAGCGGGGTGTT 841
Db CCACTGGCAGCCCGATCAGCTATTCACCTTACGCAAGTTCCTTTCGCGAGCGGGGTGCT 4235
QY 842 CAGGGGTGCTTATGACATTAATTTGTGACGAGTGCCTCCAGGATGCAACATCCA 901
Db CGGGGGCGGTTATGACATTAATTTGTGACGAGTGCCACTCCACGGATGCCACATCCA 4295
QY 902 TCTTTGGGCAATTTGGCACTGCTTTCGCAAGCAGAGACCGGGGGGAGAGTCACTGTGTC 961
Db TCTTTGGGCAATTTGGCACTGCTTTCGCAAGCAGAGACTGCGGGGGGAGAGTCTGTTGTGC 4355
QY 962 TCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCGCCCACTTAACATCGAGAGTTG 1021
Db TCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCGCCCACTCCCAACATCGAGAGTTG 4415
QY 1022 CTCTGTCCACTACGGAGAGATCCCTTTTATGGCAGGCTATTCCTCTTGAAGCAATTA 1081
Db |||||

Db 5430 GTGATAGCAGGGCAGCTGCTGTGCGCCCGGCCATTTCTACTTGAAGGCTCTCTCGG 5371
Qy 422 GAGGCTCTGCTGTGCCCCCGCAGGACATCGCGTAGGCATATTCAGAGCCGCGGTATGCA 481
Db 5370 GGGGTCCGCTGTTGTGCCCCCGCGGGCAGCCCGTGGGCATATTTAGGGCCGCGGTGTGCA 5311
Qy 482 CCCGTGAGGTGGCTAAGCGCGGTGACCTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 541
Db 5310 CCCGTGAGGTGGCTAAGCGCGGTGACCTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 5251
Qy 542 GGTCCCGGTGTTCTCAGACAACCTCTCCACCAGCAGTGCCTCCAGAGCTTACCAAGTGG 601
Db 5250 GGTCCCGGTGTTCCAGGATAACTCTCTCTCCACCAGTAGTGCCCGCAGAGCTTCCAGGTGG 5191
Qy 602 CCACCTGATGCTCCACCGCAGCGGTAAGACGACCAAGGTCCCGCGCGCATATGCGAG 661
Db 5190 CTCACCTCCATGCTCCACAGGAGCGGGCAAAAGCACCAAGGTCCCGGTGTCATATGCGAG 5131
Qy 662 CTCAGGCTCAAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 721
Db 5130 CTCAGGCTAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACAATGGGCTTTGGTG 5071
Qy 722 CTTACATGCTCAAGGCGCATGGGATGATCCTAACATCAGGACTGGGGTGAGGACAATTA 781
Db 5070 CTTACATGCTCAAGGCTCATGGGATCGATCCTAACATCAGGACCGGGTGAGGACAATTA 5011
Qy 782 CTTACTGCGAGCCGATCAGCTATTCACCTACGGCAAGTTCCTTGGCAGCGGGGTGTT 841
Db 5010 CCACCTGCGAGCCGATCAGCTACTCCACCTACGGCAAGTTCCTTGGCAGCGGGGTGCT 4951
Qy 842 CAGGGGTGCTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAACATCCA 901
Db 4950 CGGGGGCGGTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCCACATCCA 4891
Qy 902 TCTTGGSCATTTGGCACTGCTCTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 961
Db 4890 TCTTGGSCATCGGCACCTGCTCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGCTGTGTC 4831
Qy 962 TCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCCCATCTAACATCAGGAGGTTG 1021
Db 4830 TCGCCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCCCATCTAACATCAGGAGGTTG 4771
Qy 1022 CTCTGTCCACTACCGAGAGATCCCTTTTATGCAAGGCTATTCCTCTTGAAGCAATTA 1081
Db 4770 CTCTGTCCACCGAGAGATCCCTTTTATGCAAGGCTATTCCTCTTGAAGCAATTA 4711
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGAGTGCAGAGCTCGCGCAA 1141
Db 4710 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGAGTGCAGAGCTCGCGCAA 4651
Qy 1142 AACTGTGCGGTGGGGCTCAATGCGGTGCTTACTACCGCGCTTGTATGTGCTCGTCA 1201
Db 4650 AGCTGTGCGATTTGGGCATCAATGCGGTGCTTACTACCGCGCTTGTAGCTCGGTCA 4591
Qy 1202 TCCGACCACTGTGAGTGTGCTGTGCACTGAGCCCTCATGACCGGCTTTACCG 1261
Db 4590 TCCGACCACTGTGAGTGTGCTGTGCACTGAGCCCTCATGACCGGCTTTACCG 4531
Qy 1262 GCACATTCGATTCGGTGATAGACTGCAACAGTGTGTCAACCCAGACAGTGCATTCAGCC 1321
Db 4530 GCACATTCGATTCGGTGATAGACTGCAACAGTGTGTCAACCCAGACAGTGCATTCAGCC 4471
Qy 1322 TTGACCTTACCTTACCATTTGAGACATCAAGCTTCCCGAGGATGTGTCTCCCGTACTC 1381
Db 4470 TTGACCTTACCTTACCATTTGAGACATCAAGCTTCCCGAGGATGTGTCTCCCGTACTC 4411
Qy 1382 AACGTGGGGTAGGACTGGCAGAGGGAAGCAGGCACTTACACATTTGTGGCACCGGGG 1441
Db 4410 AACGTGGGGCAGACTGGCAGAGGGAAGCAGGCACTTACACATTTGTGGCACCGGGG 4351
Qy 1442 AGCGTCTCTTGGCATGTTTGACTCGGTCTGCTCTGCGAGTGTCTATGACGCGGGTGTG 1501

Db 4350 AGCGCCCTCCGCGCATGTTTCGACTCGTCCGTCTCTGTGAGTGTCTATGACGCGGCTGTG 4291
Qy 1502 CTTGGTATCAGCTTACGCGCGCGAGACACACAGTTAGGCTACGAGCATACATGAACACCC 1561
Db 4290 CTTGGTATCAGCTTACGCGCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACCC 4231
Qy 1562 CGGGACTTCCCGTGTGCCAAGACCATCTTTGAATTTTGGGAGGCGTCTTTACGGGTCTCA 1621
Db 4230 CGGGCTTCCCGTGTGCCAAGACCATCTTTGAATTTTGGGAGGCGTCTTTACAGGCTCA 4171
Qy 1622 CCACATAGACGCGCACTTCTCTATCCAGACAAGCAGAGTGGGAAAAACCTTCCCTATC 1681
Db 4170 CTCATATAGATGCCCATTTCTTATCCAGACAAGCAGAGTGGGAGAACCTTCTTACC 4111
Qy 1682 TGGTAGCTACCAAGCACCGTGTGCGCTAGAGCTCAAGCCCTCCATGGGCCAACACCTCTCTAT 1741
Db 4110 TGGTAGCTACCAAGCACCGTGTGCGCTAGGCTCAAGCCCTCCATGGGCCAACACCTCTCTAT 4051
Qy 1742 AGATGTGGAAGTCTTGTATCCGCTCTCAAGCCACCCCTCCATGGGCCAACACCTCTCTAT 1801
Db 4050 AGATGTGGAAGTCTTGTATCCGCTCAAGCCACCCCTCCATGGGCCAACACCTCTCTAT 3991
Qy 1802 ATAGACTGGCGCTGTCCAGAAATGAAGTCACTGTGACGCCACCACTCAACAGTATATCA 1861
Db 3990 ACAGACTGGCGCTGTTCAGAAATGAATCACTGTGACGCCACCACTCAACAGTATATCA 3931
Qy 1862 TGACATGTATGCTGCGCTGACCTGGAGTGTGCTCAAGTACCTGGGTGCTGTTGGCGGG 1921
Db 3930 TGACATGTATGCTGCGCGACCTGGAGTGTGCTCAAGTACCTGGGTGCTGTTGGCGGG 3871
Qy 1922 TTCTGCTCTTGGCGCGTATTTGCTATCCACAGCTTGGTGTCTATAGTAGTAGGA 1981
Db 3870 TCTGCTCTTGGCGCGTATTTGCTCTCAAGCTTGGTGTCTATAGTAGTAGGA 3811
Qy 1982 TTGCTTTGTCGAAAGCGGCAATCATACCCGACAGGAACTCTCTACCGGAGTTG 2041
Db 3810 TGGTCTTGTCCGGAAGCGGCAATCATACCTGACAGGAACTCTCTACCGAGAGTTG 3751
Qy 2042 ATGAAATGGAAGTCT 2059
Db 3750 ATGAGATGGAAGTCT 3733

RESULT 14
US-08-444-818-176
; Sequence 176, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.

```

; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (508)359-3876
;   TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 176:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9379 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 345
; OTHER INFORMATION: /note= "A heterogeneity may exist at this position which is A or G"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 351
; OTHER INFORMATION: /note= "A heterogeneity may exist at this position which is A or C"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 846
; OTHER INFORMATION: /note= "A heterogeneity may exist at this position which is T or C"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1319
; OTHER INFORMATION: /note= "A heterogeneity may exist at this position which is A or G"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2126
; OTHER INFORMATION: /note= "A heterogeneity may exist at this position which is A or C"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3659
; OTHER INFORMATION: /note= "A heterogeneity may exist at this position which is C or T"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4146
; OTHER INFORMATION: /note= "A heterogeneity may exist at this position which is C or T"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4680
; OTHER INFORMATION: /note= "A heterogeneity may exist at this position which is G or A"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9080
; OTHER INFORMATION: /note= "A heterogeneity may exist at this position which is A or G"
US-08-444-918-176

Query Match      86.6%; Score 1784.4; DB 3; Length 9379;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 1884; Conservative 4; Mismatches 170; Indels 0; Gaps 0;

QY    2   TGGCGCCCTATCAGCGGCTGTGCCAGACAGCAAGGGCCCTTTTGGGATGCATAATCACCA 61
Db     3396 TGGCGCCCTATCAGCGGCTGTGCCAGACAGCAAGGGCCCTCTTAGGGTGATATACCA 3455
DY     62  GTTTGACC GCCGGGACAAAACACAGGTGGAGGGTGAGGGTTTCAGATCGTGTCAACTGCTG 121

```

Db 4536 AGCTGTCGCAATTGGGCAATCAATGCGGTGACCTATACCGCGGTCTTGACGTCCTGCTCA 4595
Qy 1202 TCCGACACAGTGTGAGTGTGCTGCTGTCGCAACTGACGCGCTCATCACGCGCTTTACCG 1261
Db 4596 TCCGACACAGCGGATGTTGCTGCTGTCGCAACCGATGCTCATGACCGCTTATACCG 4655
Qy 1262 GCACATTCGATTCCGTTGATAGACTGCAACAGTGTGTCACCCAGACAGTGCATTCAGCC 1321
Db 4656 GCACATTCGACTCGGTGATAGACTRCAATAGTGTGTCACCCAGACAGTGCATTCAGCC 4715
Qy 1322 TTGACCTTACTCTTACCAATTGAGCAATCAGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381
Db 4716 TTGACCTTACTCTTACCAATTGAGCAATCAGCTTCCCGAGGATGCTGTCTCCCGACTC 4775
Qy 1382 AAGCTCGGGTAGGACTGGCAGAGGGAAGCAGGCATCTACAGATTGTGSCACCGGGG 1441
Db 4776 AAGCTCGGGCAGGACTGGCAGAGGGAAGCAGGCATCAACAGATTGTGSCACCGGGG 4835
Qy 1442 AGGCTCCTCTTGGCATGTTTGACTCGTCTGCTCTGCGAGTGTCTATGACGCGGTTGTG 1501
Db 4836 AGCGCCCTCCGCAATGTTTCGACTCGTCTGCTCTGCTGAGTGTCTATGACGCGGCTGTG 4895
Qy 1502 CTTGGTATAGCTTACGCCCGCGAGACACAGTTAGGCTACAGACATACATGAACACC 1561
Db 4896 CTTGGTATAGCTTACGCCCGCGAGACTACAGTTAGGCTACAGAGCGTACATGAACACC 4955
Qy 1562 CGGACTTCCCGTGTGCAAGACCATCTTGAATTTTGGGAGGCGCTTTTACGGGCTCA 1621
Db 4956 CGGCGCTTCCCGTGTGCAAGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCTCA 5015
Qy 1622 CCACATAGACGCCCACTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 1681
Db 5016 CTCTATAGATGCCCACTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTTACC 5075
Qy 1682 TGTAGGCTACCAAGCACCGTGTGCTGCTAGAGCTCAAGCCCTCCCGCTGTGGGACC 1741
Db 5076 TGTAGGCTACCAAGCACCGTGTGCTGCTAGGCTCAAGCCCTCCCGCTGTGGGACC 5135
Qy 1742 AGATGTGAAGTGTGATCCGCTCAAGCCCAAGCTTCAATGGGCAACACCTCTGCTAT 1801
Db 5136 AGATGTGAAGTGTGATCCGCTCAAGCCCAAGCTTCAATGGGCAACACCTCTGCTAT 5195
Qy 1802 ATAGACTGGCGCTGTCCAGATGAAGTCAACCTGACGCAACCCAGTCAACCAAGTATCA 1861
Db 5196 ACAGACTGGCGCTGTTCAGATGAATCAACCTGACGCAACCCAGTCAACCAATATCA 5255
Qy 1862 TGACATGATATGTCGGCTGACCTGGAGTGTCTACGAGTACTGGGTGCTGTGGCGCG 1921
Db 5256 TGACATGATATGTCGGCGACCTGGAGTGTCTACGAGTACTGGGTGCTGTGGCGCG 5315
Qy 1922 TTCTGCTGCTTTGGCGCGTATTGCTATCCACAGGCTCGGTGATAGTAGTAGGA 1981
Db 5316 TCCTGCTGCTTTGGCGCGTATTGCTATCCACAGGCTCGGTGATAGTAGTAGGAGG 5375
Qy 1982 TTCTGCTGCTTTGGCGCGTATTGCTATCCACAGGCTCGGTGATAGTAGTAGGAGG 2041
Db 5376 TCCTGCTGCTTTGGCGCGTATTGCTATCCACAGGCTCGGTGATAGTAGTAGGAGG 5435
Qy 2042 ATGAATGGAAGTGTCT 2059
Db 5436 ATGAGATGGAAGTGTCT 5453

RESULT 15

US-09-881-239-2
; Sequence 2, Application US/09881239
; Patent No. 6630298
; GENERAL INFORMATION:
; APPLICANT: CHIEN David Y.
; APPLICANT: ARANGEL, Phillip
; APPLICANT: TANDESKE, Laura
; APPLICANT: GEORGE-NASCIEMENTO, Carlos
; APPLICANT: COIT, Doris

; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY
; FILE REFERENCE: 2302-16073 / PP16073.003
; CURRENT APPLICATION NUMBER: US/09/881,239
; CURRENT FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: representative NS3/4a conformational antigen
; NAME/KEY: CDS
; LOCATION: (1)..(2058)
US-09-881-239-2

Query Match 86.4%; Score 1781.2; DB 4; Length 2058;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1885; Conservative 0; Mismatches 173; Indels 0; Gaps 0;
Qy 1 ATGGCGCCTATCACGCGCTATGCCAGCAGACAGAGGGCGCTTTGGGATGCATAATCAC 60
Db 1 ATGGCGCCTATCACGCGCTATGCCAGCAGACAGAGGGCGCTTCTAGGGTGCATAATCAC 60
Qy 61 AGCTTCCACCGCGCGGACAAAACACAGGTGGAGGTGAGGTTCCAGATCGTGTCAACTGCT 120
Db 61 AGCTTAACTTGGCGGGACAAAACCAAGTGGAGGTGAGGTTCCAGATTTGTCTCACTGCT 120
Qy 121 GCCAGACTTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGACTGTCTACCAATGAGGCC 180
Db 121 GCCCAAACTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGACTGTCTACCAAGGGCC 180
Qy 181 GGAACAGGACCAATTCGCTCACCTAAGGCTCTGTTATCCAGATGATACCAATGTGGAC 240
Db 181 GGAACAGGACCAATTCGCTCACCTAAGGCTCTGTTATCCAGATGATACCAATGTAGAC 240
Qy 241 CAAGACTCTGTAGGCTGGCGCGCTCCCAAGGTGCGGCTCATTAACACCATGCACCTTGC 300
Db 241 CAAGACTTTGTGGCTGGCGCGCTCCCAAGGTGCGGCTCATTAACACCATGCACCTTGC 300
Qy 301 GGCTCTCTCGACCTTTTACCTGTGTACAGGACAGCGCGATGTCTTCTGTGCGCCGACGG 360
Db 301 GGCTCTCTCGACCTTTTACCTGTGTACAGGACAGCGCGATGTCTTCTGTGCGCCGCG 360
Qy 361 GGTGATGAGGCGGAGCGCTTCTTCCCGCGCGCTATCTCTTACTTGAAGGCTCCTCG 420
Db 361 GGTGATGAGGCGGAGCGCTTCTTCCCGCGCGCTATCTCTTACTTGAAGGCTCCTCG 420
Qy 421 GGAGGCGCTCTGCTGTGCGCGCAGACATGCGTAGGCATATTCAGAGCGCGGTATGC 480
Db 421 GGAGGCGCTCTGCTGTGCGCGCAGACATGCGTAGGCATATTCAGAGCGCGGTATGC 480
Qy 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTTATCCCTGTGGAGAACCTTAGAGCAACCATG 540
Db 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTTATCCCTGTGGAGAACCTTAGAGCAACCATG 540
Qy 541 AGGTCCCGGTGTCTCAGACAACTCTCCCGCAGCAGTGTGCCAGAGCTACCAAGTG 600
Db 541 AGGTCCCGGTGTCTCAGACAACTCTCCCGCAGCAGTGTGCCAGAGCTACCAAGTG 600
Qy 601 GCCCACTGCTATGCCCGCGGCTAAGAGCACCAAGGTCCCGGCGCATACGCA 660
Db 601 GCTCCTCTCATGCTCCCGCAGCAGGCAAGAACCAAGGTCCCGGCTGATATGCA 660
Qy 661 GCTCAGGCTTACAAGGTGCTGTGCTCAACCCCTCCCTGTGCTCAACAATTTGGCTTTGGT 720
Db 661 GCTCAGGCTTACAAGGTGCTGTGCTCAACCCCTCCCTGTGCTCAACAATTTGGCTTTGGT 720
Qy 721 GCTTACATGTCGAAGGCTATGGGATTCCTTAACTCAGGACTGGGTGAGACAATTT 780
Db 721 GCTTACATGTCGAAGGCTATGGGATTCCTTAACTCAGGACTGGGTGAGACAATTT 780

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 15:35:02 ; Search time 6101 Seconds
(without alignments)
12858.623 Million cell updates/sec

Title: US-09-930-591-1

Perfect score: 2061

Sequence: 1 atggcgctatcacggccta.....atgaatggaagagtgtga 2061

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	46.8	2.3	509	9	CL252974
C 2	44.8	2.2	935	9	CNS006XK
C 3	44.4	2.2	834	4	BI958973
C 4	43.4	2.1	666	5	BM950570
C 5	42.8	2.1	1863	7	CV068974
C 6	42.6	2.1	633	4	BI959933
C 7	42.2	2.0	871	6	CD437613
C 8	41.8	2.0	564	2	BE291962
C 9	41.8	2.0	673	4	BI155705
C 10	41.8	2.0	935	5	BQ876256
C 11	41.8	2.0	1020	6	BY711719
C 12	41.8	2.0	1409	3	AK012576
C 13	40.8	2.0	925	9	CNS0091P
C 14	40.4	2.0	533	4	BZ277484
C 15	40.4	2.0	1749	9	CU948408
C 16	40.2	2.0	574	4	BZ208789
C 17	40.2	2.0	629	4	BG906349
C 18	40.2	2.0	684	4	BI956133
C 19	40	1.9	697	7	CN944784
C 20	39.8	1.9	624	6	CD938118
C 21	39.8	1.9	896	4	BG965973
C 22	39.6	1.9	424	5	BQ665971
C 23	39.6	1.9	429	5	BQ665888
C 24	39.6	1.9	430	5	BQ665954

C 25	39.6	1.9	574	6	CB873738
C 26	39.6	1.9	595	6	CA013559
C 27	39.6	1.9	641	6	CD861674
C 28	39.6	1.9	751	7	CF452146
C 29	39.6	1.9	848	7	CV069194
C 30	39.6	1.9	925	9	CNS0091P
C 31	39.4	1.9	913	5	BQ681158
C 32	39.4	1.9	950	9	CC999197
C 33	39.2	1.9	508	6	CA615639
C 34	39.2	1.9	753	4	CW004754
C 35	39	1.9	373	4	BJ281690
C 36	39	1.9	1146	9	CL648524
C 37	38.8	1.9	645	6	CA600316
C 38	38.8	1.9	780	9	CL944249
C 39	38.8	1.9	1041	9	CL946173
C 40	38.6	1.9	399	6	CB655390
C 41	38.6	1.9	438	6	CB927725
C 42	38.6	1.9	534	2	BF291714
C 43	38.6	1.9	552	6	CB926197
C 44	38.6	1.9	559	6	CB926189
C 45	38.6	1.9	596	6	CA032454

ALIGNMENTS

RESULT 1

LOCUS

CL252974/c

DEFINITION

ZMMBBB060101r ZMMBBB (HindIII) zea mays genomic clone

ACCESSION

CL252974

VERSION

CL252974.1 GI:41356103

KEYWORDS

GSS.

SOURCE

Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

REFERENCE

1 (bases 1 to 509)

AUTHORS

Bharti, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C., Zohovetz, V., Fuks, G., Yu, Y., Wing, R. and Messing, J.

TITLE

Sequencing of the maize genome at PCR (2003c)

JOURNAL

Unpublished (2003)

COMMENT

Contact: Bharti, A.K.

Dr. Joachim Messing's lab

The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University

190 Frelinghuysen Road, Piscataway, NJ 08854, USA

Tel: 732 445 3801

Fax: 732 445 5735

Email: bharti@waksman.rutgers.edu

Seq primer: SP6

Class: BAC ends

High quality sequence start: 73.

FEATURES

Location/Qualifiers

1..509

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="ZMMBBB060101"

/lab_host="E. coli DH10B"

/clone_lib="ZMMBBB (HindIII)"

/note="vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match

Best Local Similarity

Matches 114; Conservative

2.3%; Score 46.8; DB 9; Length 509;

50.4%; Pred. No. 0.2;

Mismatches 112; Indels 0; Gaps 0;

QY 1147 GTCCGGTGGCGCAATCGCTGCTTACTACCGCGCCTGATGTCGTCATCCG 1206


```

Db      406 GTCTGTCGTCCTTCTGTCGTCGTGTCACCTTTGTGTCGTCGTCTACTGTCTATGTC 347
QY      1207 ACCAGTGGTGACCTGTGTCGTGGCAACTGAGCCCTCATGACCGGCTTTTACCGCGAC 1266
Db      346 GCCTTGGCGTCATGTTGTGGCCATAGTGTGCGCCTCGCCGTGCCATCACCTTGCC 287
QY      1267 TTGATTCGGTGTAGACTGCAACAGCTGTGTACCCAGACAGTGCATTCAGCCTTGAC 1326
Db      286 GTTGCCACACTTCGCGGTGCGCGTCAACCGCGCCACCGCCCTTGCCTGCTGTGTGCTGCT 227
QY      1327 CTTACCTTACCATTTAGACAATCAAGCTTCCCGAGATGCTGCT 1372
Db      226 GCTGCGGTGTCGTCGCGCCATCACCTCGCCACTGCTCCACTCT 181

```

RESULT 2

CNS006XK/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14N09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL066051

VERSION
AL066051.1

KEYWORDS
GSS.

SOURCE
Drosophila melanogaster (fruit fly)

ORGANISM
Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 935)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org/TheBDGP/Drosophila>

melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mammoré in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain Y2; cn bw sp, the same strain used for the BDGP's

PI and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1..935

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR14N09"

/clone_lib="RPCI-98"

/note="end : T7"

ORIGIN

```

Query Match      2.2%; Score 44.8; DB 9; Length 935;
Best Local Similarity 30.9%; Pred. No. 0.84;
Matches 94; Conservative 64; Mismatches 146; Indels 0; Gaps 0;

QY      356 GACGGGTGATGGCAGGGGAGCTGCTTTGCGCCGGCCTATCTCTTACTTGAAGGCT 415
Db      927 GSGSSSSGSGCGSGCGSGCGSCCGSCCGSCCGSCCGSCCGSCCGSCCGSCCG 868
QY      416 CTTGCGGAGGCGCTCTCTGTGCGCCCGCAGACATGCGGTAGGCATATTCAGAGCCCGG 475
Db      867 CGCCSCGSGSCCGCGCGSSSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808
QY      476 TATGCACCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGAGCTTAGACAA 535

```

```

Db      807 CSBGCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGSSS 748
QY      536 CCATGAGGTCCCGCGGTGTTCTCAGACAACTCTCTCCCGCAGCAGTGCCTCCAGAGTACC 595
Db      747 GSGCGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 688
QY      596 AAGTGCCCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
Db      687 GSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 628
QY      656 AQCG 659
Db      627 GGCG 624

```

RESULT 3

BI956973

LOCUS

DEFINITION

(normal) Hordeum vulgare subsp. vulgare cDNA clone HVSME0006116f,

mRNA sequence.

ACCESSION

BI956973

VERSION

BI956973.1

KEYWORDS

SOURCE

ORGANISM

Hordeum vulgare subsp. vulgare

Hordeum vulgare subsp. vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Hordeum.

REFERENCE

1 (bases 1 to 834)

Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Chin, A., Begum, D.,

Frisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T.,

Simmons, J., Oates, R. and Main, D.

Development of a genetically and physically anchored EST resource

for barley genomics: Morex rachis cDNA library

Unpublished (2001)

Contact: Wing RA

Clemson University

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: wing@clemson.edu

Total hg bases = 471

Seq primer: AATTAACCTCTACTAAAGG

High quality sequence stop: 733.

FEATURES

source

1..834

/organism="Hordeum vulgare subsp. vulgare"

/mol_type="mRNA"

/cultivar="Morex"

/sub_species="vulgare"

/db_xref="taxon:112509"

/clone="HVSME0006116f"

/tissue_type="Rachis"

/lab_host="TJC121"

/clone_lib="Hordeum vulgare rachis EST library HVCDA0015

(normal)"

/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; Plants were grown at Washington State University,

Pullman, WA in a greenhouse, the rachises were excised and

frozen in liquid nitrogen (Kleinhofs lab). In the TJ Close

lab at the University of California, Riverside total RNA

was prepared, poly(A) was purified, one primary

unamplified cDNA library was made, and 1 million pfu were

in vivo excised to give phagescript SK(-) cDNA phagemids

(Chin). Phagemids were plated and picked at the Clemson

University Genomics Institute (CUGI) (Begum, Palmer,

Frisch, Atkins and Wing). Plasmid DNA preparations, DNA

sequencing and sequence analysis were performed at CUGI

(Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The

sequence has been trimmed to remove vector sequence and

```
seq primer: pix-3.
FEATURES             Location/Qualifiers
     1..666
         /organism="Mus musculus"
         /mol_type="mRNA"
         /strain="C57BL/6"
         /db_xref="taxon:10090"
         /clone="IMAGE:5687271"
         /tissue_type="whole brain"
         /dev_stage="embryo 18.5 dpc"
         /lab_host="DH10B (T1 phage resistant)"
         /clone_lib="NIH_EMAP_EH04"
     /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
     Site 2: Not I; The library was constructed according to
```

ORIGIN

(Gulliard's F/2 media).

Query Match 2.1%; Score 42.8; DB 7; Length 1863;
Best Local Similarity 48.4%; Pred. No. 3.7;
Matches 119; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 1126 GACGAGCTCGCGCAAAACTGGTTCGGTGGGGCGTCAATGCCGTGGCTTACTACCGCGC 1185

1523 GACGCCACAGCGCCATAGCGGACCCCTGCTCCATAGCGCCCTCCCCCATTTGCC 1464
 1186 CTTGATGTGTCCTATCCCGACCACTGTGTGTCGTGTGCAACTGACGCCCTTC 1245
 1463 GCGGGGTTTACCGGCTTCCCGGGCGCGCGCCAGGCGCTTCCACCGCGCGCTTC 1404
 1246 ATGACCGGCTTACCGGCGACTTCGATTGGGTATAGATGCAACAGCTGTGTACCCAG 1305
 1403 AAGCCACACCTTCCCGCGCGGCTTGGCGCGGTGCGGCCACCGCACCGCGGTTCCTTT 1344
 1306 ACAGTGCAGCTTACGCTTACCGCTTACCTTACCATTCAGACATACAGCTTCCCCAGGAT 1365
 1343 GTCGCGGCTTCCCGCCATCGCCACTTTCGCCACCGTCGCCATCGGCCACCGTCGGAT 1284
 1366 GCTGTC 1371
 1283 TCTGAC 1278
 RESULT 6
 BI959933
 LOCUS
 DEFINITION
 HVSMen0022H07f Hordeum vulgare rachis EST library HVCDNA0015
 (normal) Hordeum vulgare subsp. vulgare cDNA clone HVSMen0022H07f,
 mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Hordeum vulgare subsp. vulgare
 Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticeae; Hordeum.
 1 (bases 1 to 633)
 Wing, R., Close, T.J., Klein, H.A., Wise, R., Chin, A., Begum, D.,
 Frisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T.,
 Simmons, J., Oates, R., and Main, D.
 Development of a genetically and physically anchored EST resource
 for barley genomics: Morex rachis cDNA library
 Unpublished (2001)
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hg bases = 513
 Seq primer: AATTAAACCTTCACTAAAGGG
 High quality sequence stop: 595.
 Location/Qualifiers
 1. 633
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Morex"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clones="HVSMen0022H07f"
 /tissue_type="Rachis"
 /lab_hosts="TJC121"
 /clone_lib="Hordeum vulgare rachis EST library HVCDNA0015
 (normal)"
 /note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; Plants were grown at Washington State University,
 Pullman, WA in a greenhouse, the rachis were excised and
 frozen in liquid nitrogen (Klein, H.A. lab). In the TJ Close
 lab at the University of California, Riverside total RNA
 was prepared, poly(A) was purified, one primary
 unamplified cDNA library was made, and 1 million pfu were
 in vivo excised to give pBluescript SK(-) cDNA phagemids
 (Chin). Phagemids were plated and picked at the Clemson
 University Genomics Institute (CUGI) (Begum, Palmer,

Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
 sequencing and sequence analysis were performed at CUGI
 (Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The
 sequence has been trimmed to remove vector sequence and
 contains a minimum of 100 bases of phred value 20 or
 above. For more details on library preparation and
 sequence analysis see
 http://www.genome.clemson.edu/projects/barley. To order
 this clone see http://www.genome.clemson.edu/orders Also
 see Close TJ, Wing R, Klein, H.A., Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

Query Match 2.1%; Score 42.6; DB 4; Length 633;
 Best Local Similarity 51.3%; Pred. No. 3.1;
 Matches 99; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1096 CTCATCTTCTGCACTCAAAAGAGAGTGCAGAGCTCGCGCAAACTGGTTCGGTTG 1155
 Db 405 CTCCTCACTGTCCAGCGGACAGACTCGGGGGCCCCAACGGGACGGTTCGGCGG 464
 QY 1156 GCGGTCAATGCGGTGCTTACTACCGCGCTTGTGTGTCTATCCGACCACTGGT 1215
 Db 465 AGCATCAACAGCGCTCTCTGCTGCTCCCTCCACCGTCTCCATCTCCAGGCACACTAC 524
 QY 1216 GAGTGTGTGTGCGCACTGACCGCTCATGACCGGCTTACCGCGGACCTTCGATTG 1275
 Db 525 CAGGGCGATGCGCGCGCCACCGGCTTCCACCGGACTTTCCCGCCCAACCGCGG 584
 QY 1276 GTGATAGACTGCA 1288
 Db 585 CAGTTCGACTACA 597

RESULT 7
 CD437613
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 871)
 Lai, J., Dey, N., Kim, C.S., Bharti, A.K., Rudd, S., Mayer, K.F.,
 Larkins, B., Becraft, P., and Messing, J.
 Characterization of the maize endosperm transcriptome and its
 comparison to the rice genome
 Genome Res. 14 (10), 1932-1937 (2004)
 Contact: Lai, Jinsheng
 Dr. Joachim Messing's lab
 Waksman Institute, Rutgers University
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
 Tel: 732-445-3801
 Fax: 732-445-5735
 Email: jlai@waksman.rutgers.edu
 Seq primer: T3.
 Location/Qualifiers
 1. 871
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="W22"
 /db_xref="taxon:4577"
 /tissue_type="Endosperm of 7-23DAP"
 /clone_lib="Endosperm 5"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI"

CD437613 871 bp mRNA linear EST 03-JUN-2003
 ELOIN0503A05.b Endosperm_5 Zea mays cDNA, mRNA sequence.
 CD437613
 CD437613.1 GI:31353256
 EST.
 Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 871)
 Lai, J., Dey, N., Kim, C.S., Bharti, A.K., Rudd, S., Mayer, K.F.,
 Larkins, B., Becraft, P., and Messing, J.
 Characterization of the maize endosperm transcriptome and its
 comparison to the rice genome
 Genome Res. 14 (10), 1932-1937 (2004)
 Contact: Lai, Jinsheng
 Dr. Joachim Messing's lab
 Waksman Institute, Rutgers University
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
 Tel: 732-445-3801
 Fax: 732-445-5735
 Email: jlai@waksman.rutgers.edu
 Seq primer: T3.
 Location/Qualifiers
 1. 871
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="W22"
 /db_xref="taxon:4577"
 /tissue_type="Endosperm of 7-23DAP"
 /clone_lib="Endosperm 5"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI"

FEATURES
 source
 ORIGIN


```

RESULT 10
BQ876256/c
LOCUS
DEFINITION BQ876256 935 bp mRNA linear EST 16-AUG-2002
5', mRNA sequence.
ACCESSION BQ876256
KEYWORDS BQ876256.1 GI:22268262
SOURCE EST.
ORGANISM Mus musculus (house mouse)
Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 935)
NIH-MGC http://mgi.nci.nih.gov/
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cga@rs-rcmail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: ResGen, Invitrogen Corp
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M13875 row: m column: 14
High quality sequence stop: 761.
FEATURES
Location/Qualifiers
1..935
/mol_type="mRNA"
/db_xref="taxon:10090"
/lab_host="IMAGE:6389941"
/clone_lib="NIH MGC 129"
/note="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;
Site1: EcoRV; Site2: NotI; Cloned unidirectionally.
Primer: Oligo dt. Average insert size 2.2 kb. Constructed
by ResGen, Invitrogen Corp. Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 2.0%; Score 41.8; DB 5; Length 935;
Best Local Similarity 53.3%; Pred. No. 5.8;
Matches 88; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
687 CAACCCCTCGTGTGCTGCAACATGGCTTTGGTGTTCATCTCCAGGCCCATGGAT 746
|||||
548 CAAGGCCACCTGTGCTGGAGCCATGGCTCTGCAGTGTCTCTGCTCAGAGCCACAGGCT 489
|||||
747 TGATCTTAACATCAGGACTGGGTGAGGACAACTACTCTGGCAGCCCGATCATGTTTC 806
|||||
488 GGACACAAAGGGGAGCTCTCTCGGTAGGGCAAGGTCACAGGAGCGCTGGGCTAGTGCC 429
|||||
807 CACTACGGCAAGTCTCTTCCGACGCGGTGTTCAGGGGGTGC 851
|||||
428 CGGCTCTGGAGCAGCTGCAGCCGCGGTGGAGTTCAGCTGAGGC 384
|||||

RESULT 11
BY711719/c
LOCUS
DEFINITION BY711719 RIKEN full-length enriched, 11 days embryo whole body Mus
musculus cDNA clone 2700087H15 5', mRNA sequence.
ACCESSION BY711719
KEYWORDS BY711719.1 GI:27122976
SOURCE EST.
ORGANISM Mus musculus (house mouse)
Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 1020)

```

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batilov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Malcais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomica, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kitahara, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-2 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Nomura, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES

source

Location/Qualifiers
1..1020
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2700087H15"

5

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2005, 01:27:56 ; Search time 6776 Seconds
(without alignments)
4905.590 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPTAYAQTRGLGCIIT.....PALIPREVLRYEPEMEEC 686

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09930591/runat_25022005_143612_25675/app_query.fasta_1.839
-DB=GenEmbl -QWTF=fastcap -SUFFIX=rgc -MINWATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09930591.OCGN 1.1 5600 @runat_25022005_143612_25675 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sv.*

13: gb_un.*

14: gb_vi.*

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	3618	100.0	2061	CQ826998 Sequence
2	3618	100.0	2061	AX441176 Sequence
3	3618	100.0	2061	AX467113 Sequence
4	3618	100.0	2078	CQ827032 Sequence

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	3579	98.9	8791	14	AY615798	AY615798 Hepatitis I06434 Sequence 48
6	3574	98.8	5360	6	I06434	I09328 Sequence 8
7	3574	98.8	5360	6	I09328	I06440 Sequence 54
8	3574	98.8	6785	6	I09329	I09329 Sequence 10
9	3574	98.8	6785	6	AR118696	AR118696 Sequence
10	3574	98.8	7310	6	I09331	I09331 Sequence 15
11	3574	98.8	7310	6	I09331	M32084 Hepatitis C I08294 Sequence 1
12	3574	98.8	9185	6	I08294	BD091382 HCV culti AR166930 Sequence
13	3574	98.8	9185	6	BD091382	AR166930 Sequence
14	3574	98.8	9185	6	BD091382	AR166930 Sequence
15	3574	98.8	9379	6	AR166930	AR166930 Sequence
16	3574	98.8	9379	6	AR166930	AR166930 Sequence
17	3574	98.8	9401	6	AR166930	AR166930 Sequence
18	3574	98.8	9401	6	AR166930	E66593 Hepatitis C I08294 Sequence 1
19	3574	98.8	9401	6	E66593	AR176483 Sequence
20	3574	98.8	9401	6	E66593	E66593 Hepatitis C I08294 Sequence 1
21	3574	98.8	9401	6	E66593	AR176483 Sequence
22	3574	98.8	9401	6	E66593	E66593 Hepatitis C I08294 Sequence 1
23	3574	98.8	9401	6	E66593	AR176483 Sequence
24	3574	98.8	9401	6	E66593	E66593 Hepatitis C I08294 Sequence 1
25	3574	98.8	9401	6	E66593	AR176483 Sequence
26	3571	98.7	6785	6	AR118692	AR118692 Sequence
27	3570	98.7	8316	6	AR118703	AR118703 Sequence
28	3570	98.7	9185	6	AR118722	AR118722 Sequence
29	3570	98.7	9185	6	AR118722	AR118722 Sequence
30	3570	98.7	9618	14	AF271632	AF271632 Hepatitis AX164584 Sequence
31	3568	98.6	6299	6	AX164584	AX164584 Sequence
32	3565	98.5	2058	6	AR404933	AR404933 Sequence
33	3565	98.5	2058	6	AR408362	AR408362 Sequence
34	3565	98.5	2058	6	AX395309	AX395309 Sequence
35	3565	98.5	2058	6	AX454818	AX454818 Sequence
36	3565	98.5	8987	6	AR118728	AR118728 Sequence
37	3565	98.5	9609	12	AF387808	AF387808 Synthetic AR18586 Sequence
38	3555	98.3	5360	6	AR118686	AR118686 Sequence
39	3551	98.1	9646	6	AR110828	AR110828 Sequence
40	3551	98.1	9646	6	BD069982	BD069982 Functiona BD069982 Hepatitis AF003606 Hepatitis
41	3551	98.1	9646	14	AF009606	AF009606 Hepatitis AF003606 Hepatitis
42	3551	98.1	12980	6	AR110831	AR110831 Sequence
43	3551	98.1	12980	6	BD069985	BD069985 Functiona BD069985 Hepatitis
44	3550	98.1	9395	14	AF511950	AF511950 Hepatitis E08263 GRNA of Hep
45	3550	98.1	9502	6	E08263	E08263 GRNA of Hep

ALIGNMENTS

RESULT 1	CQ826998	Sequence 1 from Patent WO2004048402.	2061 bp	DNA	linear	PAT 29-JUN-2004
LOCUS	CQ826998	Sequence 1 from Patent WO2004048402.	2061 bp	DNA	linear	PAT 29-JUN-2004
DEFINITION	CQ826998	Sequence 1 from Patent WO2004048402.	2061 bp	DNA	linear	PAT 29-JUN-2004
ACCESSION	CQ826998	Sequence 1 from Patent WO2004048402.	2061 bp	DNA	linear	PAT 29-JUN-2004
VERSION	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
KEYWORDS	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
SOURCE	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
ORGANISM	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
REFERENCE	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
AUTHORS	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
TITLE	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
JOURNAL	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
FEATURES	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
source	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
Salberg, M.	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
A hepatitis c virus codon optimized non-structural ns3/4a fusion	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
Gene	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
Patent: WO 2004048402-A 1 10-JUN-2004;	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
TRIPEP AB (SE)	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
Location/Qualifiers	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
1..2061	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
/organism="synthetic construct"	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
/mol_type="unassigned DNA"	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
/db_xref="taxon:32630"	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004
/note="Hepatitis C virus NS3/4A coding region"	CQ826998.1	GI:49455655	2061 bp	DNA	linear	PAT 29-JUN-2004

ORIGIN	Alignment Scores:	7.16e-203	Length:	2061
	Pred. No.:	3618.00	Matches:	686
	Score:	100.00%	Conservative:	0

Best Local Similarity: 100.00%		Mismatches: 0
Query Match: 100.00%		Indels: 0
DB: 6		Gaps: 0
US-09-930-591-2 (1-686) x CQ826998 (1-2061)		
QY	1 MetAlaProIleThrAlaValAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20	341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLeuAlaIleProLeuGluAlaIle 360
DB	1 ATGGCGCCTATCACGGCCTATGCCAGCAGACAAAGGGGCTTTGGATGCATAATCAC 60	1021 GCTCTGTCACATACCGGAGAGATCCCTTTATGGCAAGGCTATTCCCTTGAAGCAATT 1080
QY	21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40	361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
DB	61 AGCTTGACCGCGCGGACAAAACAGGTGGAGGTGAGGTTCCAGATCGTGCACTGCT 120	1081 AAGGGGGGAGAGACATCTCATCTCTGCCACTCAAGAAGAAGTGCAGAGCTCGCGCA 1140
QY	41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60	381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB	121 GCCACACTTCTTGGCACTGCATTAACGGGGTGTGTGGACTGTCTACCATGAGCC 180	1141 AAATCGTCCGCTTGGCGCTCAATGCCGCTGCTTACTACCGCGCTTGTATGTCTCGGC 1200
QY	61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80	401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB	181 GGAACAAGGACCATTTGGTCACCTAAGGTCTCTGTTATCCAGATGTACACCAATGTG 240	1201 ATCCCGACCACTGGTGTGCTGCTGGCAACTGACCGCTCATGCCGCTTATACC 1260
QY	81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100	421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB	241 CAAGACCTCGTAGCTGGCGCTGCCCAAGGTGCCGCTCATTAACACCATCATCTGC 300	1261 GCGCACTTCGATTCGGTGATAGACTGCAACACGTTGTCCACCCAGACAGTCGACTTCAG 1320
QY	101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120	441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB	301 GGCTCTCTGGACCTTTACCTGGTGCAGAGGACGCGGATGTCATTTCTGTGCGCCGAC 360	1321 CTGACCCCTACCTTACCATTCAGCAATCAGCTTCCCAGGATGCTGTCTCCCGTACT 1380
QY	121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140	461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB	361 GGTGTGACGAGGGCAGCTCTTTCCGCCCGCCTATCTTACTTTGAAGGCTCCTCG 420	1381 CAACTCGGGGTAGGACTGCGCAGAGGAAGCCAGGCATCTACAGATTTGTGGCACC 1440
QY	141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160	481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB	421 GGAGGCGCTCTGTGTGCCCGCAGGACATGCGGTAGGCATATTCAGAGCCGCGGTATGC 480	1441 GAGCGTCTTCTGGCATGTTGACTCGTCTGCTCTGCGAGTGTATGACGCGGTTGT 1500
QY	161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180	501 AlaTrpTrpGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB	481 ACCGTGGAGTGGCTAAGCGGTGGACTTCCATCCCGTAGAGCTTAGAGACACCATG 540	1501 GCTTGTGTATGAGCTTACGCCCGCGAGACCAGATTAGCTACGACATACATGAACACC 1560
QY	181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200	521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
DB	541 AGTCCCCCGTGTCTCAGACAACCTCTCTCCACCAGCAGCTGCCAGAGCTACCAAGTG 600	1561 CCGGGACTTCCGTGTGCCAAGACCATCTTGAATTTGGAGGGCGCTTTTACGGGTCTC 1620
QY	201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220	541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB	601 GCCACCTGTCATGCTCCACCGCAGCGGTAAAGACACCAGGTCCCGCGCATACGCA 660	1621 ACCCATAGACGCCCACTTCTTATCCAGACAAAGCAGAGTGGGAAAACCTTCCCTAT 1680
QY	221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240	561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
DB	661 GCTCAGGGCTACAAAGGTGCTGGTGCTCAACCCCTCGTGTCTGCAACAATGGGCTTGGT 720	1681 CTGGTAGCGTACCACGCCACCGTGTGCGTGTAGAGCTCAAGCCCTCCCCCGTCTGGGAC 1740
QY	241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260	581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB	721 GCTTACATGTCCAAAGGCCCATGGATTGATCTTAACATCAGACTGGGCTGAGACAAATT 780	1741 CAGATGTGGAGTGTCTTGATCCGTCTCAAGCCACCCTCCATGGGCCAACACCTCTGCTA 1800
QY	261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280	601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTrpIle 620
DB	781 ACTACTGGCAGCCGATCAGCTATTCACCTACGGCAAGTTCCTTGGCCGCGGGGTGT 840	1801 TATAGACTGGCGCTCTCCAGAATGAAGTCACTCCCTGACGCACTACCAAGTAGTATC 1860
QY	281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300	621 MetThrCysMetSerAlaAspLeuGluValThrSerThrTrpValLeuValGlyGly 640
DB	841 TCAGGGGGTCTTATGACATAATAATTGTGACAGAGTGCACCTCCAGGATGCAACATCC 900	1861 ATGACATGATGTGCGCTACCTTGGAGGTGCTCAGGAGTACCTGGGTCTCGTGTGGCGC 1920
QY	301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320	641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB	901 ATCTTGGGCATTTGGCACTGTCTTGTACCAAGACGAGACCGCGGGGCGAGACTGTG 960	1921 GTTCTGGCTGCTTTGGCCGCTATTGCCCTATCCACAGGCTGCGTGTGCTAGTAGTAGG 1980
QY	321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340	661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB	961 CTGCGCACCGCTACCCCTCGGGCTCGTCTCACTGTGCCCATCTTAACATCGAGGAGTT 1020	1981 ATTGTCTTGTCCGAAAAGCCGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTC 2040
QY		681 AspGluMetGluGluCys 686
DB		2041 GATGAAATGGAAGAGTGC 2058

```
RESULT 2
AX441176          AX441176          2061 bp      DNA      linear      PAT 28-JUN-2002
LOCUS             Sequence 16 from Patent WO0213855.
DEFINITION
ACCESSION         AX441176
VERSION           AX441176.1  GI:21665758
KEYWORDS          synthetic construct
SOURCE            other sequences; artificial sequences.
ORGANISM          1
REFERENCE         Sallberg, M. and Hultgren, C.
AUTHORS           Vaccines containing ribavirin and methods of use thereof
TITLE             Patent: WO 0213855-A 16 21-FEB-2002;
JOURNAL           TRIPEP AB (SE)
FEATURES          Location/Qualifiers
source            1..2061
                  /organism="synthetic construct"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:32630"
                  /note="Hepatitis C virus NS3/4A coding region"
ORIGIN
Alignment Scores:
Pred. No.:        7,16e-203      Length:      2061
Score:            3618.00      Matches:    686
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:     0
DB:               6           Gaps:       0

US-09-930-591-2 (1-686) x AX441176 (1-2061)

Qy      1  MetAlaProIleThrAlaTyAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20
Db      1  ATGGCGCCTATACGGCCATATGCCAGCAGACAAAGGGGCCCTTTGGGATGCATATAC 60

Qy      21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThra 40
Db      61 AGCTTGACCGCCGGGACAAAACACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCT 120

Qy      41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThyHisGlyAla 60
Db      121 GCCCAGACTTTCCTTGGCAACCTGCATTAAACGGGGTGTGTGACTGTCTACCATGGAGCC 180

Qy      61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db      181 GGAAACAGGACCATGTGCTACCTAAGGGTCTGTATTCCAGATGTACCAATGTGGAC 240

Qy      81 GlnAspLeuValGlyTrpProAlaProGlnGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db      241 CAAGACCTCGTAGCTGGCCCGCTCCCCAAGGTGCGCGCTCATTAACACCATGCATTGC 300

Qy      101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db      301 GGCTCCTCGGACCTTTTACCTGTCACGAGGACGCCGATGTCAATTCCTGTGCGCGCAGCG 360

Qy      121 GlvAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db      361 GGTGATGCGAGGGCAGCGCTCTTTCGCCCGCGCTATCTCTTAAAGGCTCCTCG 420

Qy      141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db      421 GGAGGCCCTCTGCTGTGCCCCGCGAGGACATGCCGTAGGCATATTACAGAGCCGCGGTATGC 480

Qy      161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet: 180
Db      481 ACCCGTGGAGTGGTGAAGCGGTGACTTCATCCCGGTAGAGAGCTTAGACACAACCATG 540

Qy      181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db      541 AGGTCCCGCGGTGTTCTCAGAACCTCTCTCCCAACAGCAGCAGTGGCCCAAGTACCAAGTG 600

Qy      201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db      601 GCCCACCTGCTATGCTCCACCGGAGGGTAAGAGCACCAAGGTCCCGCGCGCATACGCA 660

Qy      221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db      661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCCGTTCCTGCAACAAATGGGCTTTGGT 720

Qy      241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db      721 GCTTACATGTCCAAGGCCCATGGATTGATCTTAACATCAGACTGGGGTCAGGACAATT 780

Qy      261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
Db      781 ACTACTGGCAGCCCGCATCAGTATTCCACCTACGCGAAGTTCTTGGCAGCGGGGTGT 840

Qy      281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db      841 TCAGGGGTGCTTATGACATAAATAATTTGTACAGGTGCCACTCCACGATGCAACATCC 900

Qy      301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db      901 ATCTTGGGCATTGGCACTGCTTTCACCAAGCAGACCGCGGGGGGAGACTGACTGTG 960

Qy      321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db      961 CTCGCCACCGCTACCCCTCCGGGTCGGTCACTGTGCCCATCTCTAAATCGAGAGGTT 1020

Qy      341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db      1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCTCTTGAAGCAAT 1080

Qy      361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db      1081 AAGGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGGCAGAGCTGCCGCA 1140

Qy      381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db      1141 AAATCGTTCGGTGGGGCGTCAATCGCGTGGCTTACTACCGCGGCTTGTATGTGTCGTC 1200

Qy      401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      1201 ATCCCGACCATGTGGTACGTTGTCGTGTCGCAACTCAGCGCCCTCATGACCGGCTTTACC 1260

Qy      421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db      1261 GCGCACTTCGATTCGGTGTAGACTGCACACGCTGTGTACCCAGACAGCTCGACTTCAGC 1320

Qy      441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db      1321 CTTGACCTTACCTTCACCATTGAGACAATCAGCCTTCCCAGGATGCTGTCTCCGCTACT 1380

Qy      461 GlnArgArgGlyArgThrGlyArgGlyAspProGlyIleTyrArgPheValAlaProGly 480
Db      1381 CAACGTCGGGGTAGGACTGGCAGAGGAAGCCAGGCATCTACAGATTGTGGCACCCGGGG 1440

Qy      481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db      1441 GAGCGCTCTTCTGGCATGTTTGTACTCGTCTGTCTCTGCGAGTGTATAGACCGGGTGT 1500

Qy      501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db      1501 GCTTGGTATGAGCTTACGCCCGCGGAGACACACAGCTTAGGCTACGAGCATACATAACACC 1560

Qy      521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db      1561 CCGGGACTTCCCGTGTGCCAAGACCATCTTCAATTTTGGGAGGGCGCTTTTACGGGTCTC 1620

Qy      541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db      1621 ACCCAGATAGACGCCCATCTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTTAT 1680

Qy      561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
```

Db	1681	CTGGTAGCGTACCAAGCCACCGGTGGCTAGAGCTCAAGCCCTCCCGCGTGGGAC	1740	
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600	
Db	1741	CAGATGTGGAAGTGTGTATCCGTCTCAAGCCACCGCTCCATGGGCAACACCTCTGCTA	1800	
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620	
Db	1801	TATAGACTGGCGCGTGTCCAGAAATGAAGTCAACCTCAGCGCACCGATGATATC	1860	
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640	
Db	1861	ATGACATGATGTTCGGCTGACCTGGAGTGTCTCAGGATACCTGGGTGCTCGTTGGCGGC	1920	
Qy	641	ValLeuAlaAlaLeuAlaLafyrCysLeuSerThrGlyCysValValIleValGlyArg	660	
Db	1921	GTTCTGGCTCTTTGGCGCGTATTGCTATTCACAGGCTGCGTGGTCAAGTAGGTAGG	1980	
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680	
Db	1981	ATTGTCTTGTCCGGAAGCCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTC	2040	
Qy	681	AspGluMetGluGluCys 686		
Db	2041	GATGAATGGAGAGTGC 2058		
RESULT 3				
LOCUS	AX467113	2061 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0214362.			
ACCESSION	AX467113			
VERSION	AX467113.1	GI:21900425		
KEYWORDS	synthetic construct			
SOURCE	synthetic construct			
ORGANISM	other sequences; artificial sequences.			
REFERENCE	1			
AUTHORS	Sallberg, M.			
TITLE	A hepatitis c virus non-structural ns3/4a fusion gene			
JOURNAL	Patent: WO 0214362-A 1 21-FEB-2002;			
FEATURES	TRIPEP AB (SE)			
source	Location/Qualifiers			
	1..2061			
	/organism="synthetic construct"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:32630"			
	/note="Hepatitis C virus NS3/4A coding region"			
ORIGIN				
Alignment Scores:				
Pred. No.:	7.15e-203	Length:	2061	
Score:	3619.00	Matches:	686	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	6	Gaps:	0	
US-09-930-591-2 (1-686) x AX467113 (1-2061)				
Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20	
Db	1	ATGGCGCTATCAGCGCTATGCCAGCAGACAAGGGCGCTTTTGGGATGCATAATCACC	60	
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla	40	
Db	61	AGCTTGACCGCGCGGCAAAACACAGTGGAGGGTGAAGTTCAGATCGTGTCAACTGCT	120	
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyVala	60	
Db	121	GCCAGACTTTCTTGGCAACTGATTAACGGGGTGTGTCGACTGTCTACATGGAGCC	180	
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValasp	80	

```
Db 1261 GCGGACTTCGATTCGGTGATAGACTGCAACACGCTGTGTACCCAGACAGTCGACTTCAGC 1320
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTTGACCCCTACCTTCACCATGAGACATACGCTTCCCGAGGATGCTGTCTCCCGTACT 1380
Qy 461 GlnArgArgGlyArgThrGlyArgGlyProGlyIleTyrArgPheValAlaProGly 480
Db 1391 CAACGTCCGGGTAGACTGGCAGAGGGAAGCCAGCATCTACAGATTGTGGCACCCGGG 1440
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 1441 GAGCGTCTCTTGGCACTTTGACTCTGCTCTCTCTCGAGTGTATGACGCGGTTGT 1500
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 1501 GCTTGGTATGAGCTTACGCCCGCCGAGACACAGTAGCTAGCTACGAGCATACATGAACACC 1560
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 1561 CCGGAGCTCCCGGTGTGCCAAGACCATCTTGAATTTTGGGAGGCGCTTTTACCGGCTC 1620
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 1621 ACCACATAGACGCCACTTCTATCCACAAAGCAGAGTGGGAAAACCTTCCCTAT 1680
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 1681 CTGTAGCTTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTGTGGGAC 1740
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGGAAAGTGTGTATCGCTCAAGCCACCTCCATGGGCCAACACCTCTGCTA 1800
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TATAGACTGGCGCTGTCCAGAAAGTCACTACCTGACGCCACCCAGTCACCAAGTATATC 1860
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATGACATGTATGTCCGCTGACCTGGAGTCTCAGAGTACTCTGGTGCTCTGTGGCGGC 1920
Qy 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTTCTGGCTGTTTGGCGCGTATTGCTATCCACAGGCTCGGTGGTCATAGTAGTAGG 1980
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 ATTGTCTGTCCGAAAGCCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTC 2040
Qy 681 AspGluMetGluGluCys 686
Db 2041 GATGAAATGGAAAGTGC 2058
```

```
RESULT 4
CQ827032
LOCUS CQ827032 2078 bp DNA linear PAT 29-JUN-2004
DEFINITION Sequence 35 from Patent WO2004048402.
ACCESSION CQ827032
VERSION CQ827032.1 GI:49455663
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Sallberg,M.
TITLE A hepatitis c virus codon optimized non-structural ns3/4a fusion
gene
JOURNAL Patent: WO 2004048402-A 35 10-JUN-2004;
AUTHORS TRIPE AB (SE)
FEATURES
location/Qualifiers
source 1..2078
/organism="synthetic construct"
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="codon optimized hepatitis C virus NS3/4A coding region"
12..2072
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAG44601.1"
/db_xref="GI:49455664"
/translation="MAPITAAQOTRGLLGCIITSLTRDKNOVEGEVOIVSTAATOF
LAINCGVCTVTVHAGACTRTASPKGPVIOMYTNVDOLVGMPPAQGRASPTCTCGS
SDLYLTRHADVTPRRRGDRGSLSPRISYLSKSGSGPLLPAGHAGVIFRAAVC
TRGAKAVDFIPVESLETTMRSPVSDNSPPAPVPQYVAHLHAPTSKSTKVPAA
YAGGKYVLNFSVAATMGFMYSKAHGDIPNRTGVTQVIRITTSITTSYTGKFLA
DGGSGGAYDIIICDECHSDATISLIGITVLDQATAGARTLVLATATPPGSVTVPH
PNTEVALSTTIGTIPFYKAIPLAIEAKGRHLIFCHSKKCDLAAKLVALGNVAV
YRGDVSIVTSGDVVVDALMTGFTGPDSDVIDCNTCVTQTVDFSLDPTFTLETI
TLQDASVTRQRRGRTGKPGIYRFVAPGERSGDMFSSVLCBQDAGCAWYBLTPA
ETTVRLRYNNTFGLPVCQDHLFEWGFGLTHIDAHFLSQTKSGENLPLVAYQA
TVCAQAQAPPSPDMWKILRLKPTLHGFTPLLYRLGAVQNEVTLTHPTVTKIMTCM
SADLEVVTSTWLVGVGVAALAAAYCLSTGCVIVGRIVLSGKPAIIPDREVLRYEFDE
MEEC"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 7,23e-203 Length: 2078
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
```

US-09-930-591-2 (1-686) x CQ827032 (1-2078)

```
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 12 ATGCCCCCATACCGCTACCGCCAGACACCGCGCGCTGTGGGTGCATCATCACC 71
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 72 AGCTGACCGCGCCGCAAGAACACAGGTGAGGGCGAGGTGCAGATCGTGAGCACCGCC 131
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 132 GCCAGACCTTCTTGGCCACCTGCATCAACGGCGTGTGCTGACCGTACACCGCGCC 191
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 192 GGCACCGCCACCATCGCCAGCCCGCCCGTGCATCCAGATGTACACCAACGCTGGAC 251
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 252 CAGGACCTGTGGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 311
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 312 GGCAGACGACCTGTACCTGTGTGACCGCCGACCGCGCGCGCGCGCGCGCGCGCGCG 371
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 372 GGCAGCGCGCGCGCGCGCTGTGAGCCCGCCCGCCCATCAGTACCTACCTGAAGGGCAGCAGC 431
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 432 GGCAGCGCGCGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 491
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
Db 492 ACCGCGCGCGCTGGCCAAAGCGCGTGGACTTATCCCCGTGGAGACCTTCCCGCGCGCG 551
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
```


Db 552 CGAGCCCCGCTGTTACGCGCAACAGCAGCCCCCGCGCTGCCGCCAGACTACCGGTG 611
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 612 GCCACCTGACGCCCCCACCAGCGAGCGGACAGACACCAAGGTGCCCGCGCTACGCC 671
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 672 GCCAGGGCTACAAAGGTGCTGGTGTGAACCCCGAGCGTGGCCGCCACCATGGCTTCGGC 731
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 732 GCCTACATGACGAAGGCCACCGCATCGACCCCAACATCCGCGCGCGTGGCCACCATC 791
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 792 ACCACGGGAGCCCATCATCTACAGCACTTACGGCAGTTCCTGGCCGACGGCGGTGC 851
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 852 AGCGGGCGGCGCTACGACATCATCTCGAGCGAGTGCACAGCAGCCGACGCCACCGAC 911
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 912 ATCTGGGCACTCGGACCGCTGCTGGACCGAGCGGAGACCGCGCGCGCTGACCGGTG 971
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 972 CTGGCCACCGCCACCCCGCGAGCGTACCGTGGCCCGCCCGCCCAACATCGAGAGGTG 1031
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1032 GCCCTGAGCACACCGCGGAGATCCCTTCTACGCAAGGCCATCCCTCGGAGGCCATC 1091
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 1092 AAGGGCGGCGCCACCTGTATCTTCTGCCACAGCAAGAAGTGCAGCAGCTGGCGGCC 1151
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 1152 AAGCTGTGGCTTGGCGGTGAACCGCGTGTACTTACCGCGCGCTGGACGTGAGCGGTG 1211
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 1212 ATCCCCACACGCGCGAGTGGTGGTGCCACCGCGCTGATGACCGGCTTCAAC 1271
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 1272 GCGGACTTCGACAGCGTGTGATCGACTGCAACACCTGCGTGACCCAGACCGGTGACG 1331
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1332 CTGGACCCCACTTACCATCGAGACCATCACCTGCGCCAGACCGCGCGGTGACCGCAC 1391
QY 461 GlnArgArgGlyArgThrGlyArgGlyProGlyIleTyrArgPheValAlaProGly 480
Db 1392 CAGCGCGCGCGCGACCGCGCGCGGCAAGCCCGCATCTACCGCTTCTGGTGGCGCCGCG 1451
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 1452 GAGCGCCCGACGCGCATGTTTCGACAGCAGCGGTGCTGTGCGAGTGTACGACGCGCGGTG 1511
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 1512 GCCTGGTACGAGCTGACCCCGCGGAGACCATCCGCTGGCGCTGCGCGCTCATGAAACAC 1571
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 1572 CCGCGGCTGCGCGGTGCGCAGGACCATCTGAGTTCCTGGAGGGCGGTTCACCGCGCTG 1631
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 1632 ACCACATCGACGCCCACTTCTGAGCCAGACCAAGCAGCGCGCGGAGAACCTTGGCCCTAC 1691

QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 1692 CTGGTGGCTACGAGGCCACCGTGTGCGCCCGCGCCAGGCCCGCCCGCCAGCTGGAC 1751
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1752 CAGATGTGAAGTGCCTGATCCGCTGAAGCCACCTGACGCGCCCGCCCGCTGCTG 1811
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1812 TACCGCTGGCGCGCGTGTGAGAACAGAGGTGACCTGACCCACCGCTGACCAAGTACATC 1871
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 1872 ATGACCTGATGAGCGCGCGACCTGGAGGTGGTGACCGACCGCTGGTGGTGGCGGC 1931
QY 641 ValLeuAlaAlaLeuAlaIleTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1932 GTGCTGGCGCGCTGGCGCTACTGCTGAGCAGCGCGTGGTGGTGGTGGCGCGC 1991
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1992 ATGCTGCTGAGCGCGCAAGCCCGCATCTCCCGACCGAGGTGCTGTACCGGAGTTC 2051
QY 681 AspGluMetGluGluCys 686
Db 2052 GACGAGATGGAGGAGTGC 2069

RESULT 5
AY615798
LOCUS
DEFINITION
ACCESSION
VERSION
SOURCE
KEYWORDS
ORGANISM

AY615798 8791 bp RNA linear VRL 14-JUN-2004
Hepatitis C virus isolate HCV.TWB.1 polyprotein gene, partial cds.
AY615798
AY615798.1 GI:48479029
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1 (bases 1 to 8791)
Brann,T.W., Kottlilil,S., Polis,M. and Imamichi,T.
Identification of mutations associated with interferon resistance
in HCV and HIV co-infected patients
Unpublished
2 (bases 1 to 8791)
Brann,T.W., Kottlilil,S., Polis,M. and Imamichi,T.
Direct Submission
Submitted (03-MAY-2004) LHR/CSP, SAIC-Frederick, Inc, Building 550,
Room 126, 1050 Boyles Street, Frederick, MD 21702, USA
Location/Qualifiers
1..8791
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="HCV1a"
/isolate="HCV.TWB.1"
/db_xref="taxon:11103"
/country="USA"
67..>8791
/codon_start=1
/product="polyprotein"
/protein_id="AAT44836.1"
/db_xref="GI:48479030"

CDS

FEATURES
source

PPCVIGGVGNNTLHCPTDCFRKHPEATYSRCSGSPWITPRCLVDYPRMLWHYPTCTNY
TLFKRMVTVGVHRELEAACNWTGRGCDLDRDRSESLPLLTITTOQLPFCSSMTLL
PALTTGLHLHGNIDVGYDYGVSISVMAIKWEYVILLFLLADARICSLCSFWMLL
ISOAEALENILVINAASLAGHLASFLVFCFAWYLKWKVPCMAAYAFPGMWPLLL
LLALLPORAYALDTEMAASCGVVLVGLMALTLSPHYKRYISWCLWMLVFLTRAEL
LOVWVPLPLNVGRGADATLLMCVHPALVFDITKLALLAALGPLMTLOTSLKVPYVR
QVGLIRICALARKAGAGHYVQMAIIKAGALGTIYINHLTFLRWANGLDLAVVE
PVFSQMETKUTIWAGDAAACGDIINGLVPVSARRREILLGPADGMASKGWRLLAPIT
AYAQTQRLGLLCIITSYTRDRKNQVEGEIQVSTAAQTFLATCINGVCMVTVYHGAGR
TIASPGKGVIMYNTNIDQLVGMWPAQARSITPCTCGSSDLVLTVRHADYIPVRGR
DSRGLSLSPRIISYKSSGGPGLCPAGHAGVIFRAAVCTRGVAKAVDFIPEVENLETT
MRSFVETDNSPPRAVPOSFVAHLHAPTGSCKTKVPAAYAAQYKVLNPSVAALT
GFGAYMSKAHGVDPNIRGVRITITGSPITISTYKFLADGCGCGGAYDIIICDECHS
TDAISLIGTVLQDAETAGARLVVLAATAPGSITVPHPNIEBALVSTGTEIFPYGK
ALPEALTKGGHLCIFCSKKCKDELAALVAGINAVAYRGLDVSIVPTSGDVVVVA
TDALMTGTFDSVIDCNTVQTFDSLPDFTFTIETTLTPQDAVSTQRRGRTGRG
KPGYRVFVAPCERPSGMSVCEVYDAGCAWYELTPAETTVRLRAVMTNPGLPVQC
DHLRFEGVFTGLHIDAHFLOSOTKSGENLPYLVAOATVCARAQAPPPSWDQWKC
LIRKPTLHGFTPLLYRLGAVONEVTLTHPVTKYIMTMSADLEVTSTVWLVGCVLA
ALAAVCUSTGCWVIIGRVLSGKPAVIDREVLVREFDEMEECQHVPIYIQQMMLAE
QFKQALGLLQTAGROAEVIAVQVTNNQKLEAFWAKHMNFISGIQVLAGLSTLPGN
PAIALMAFTAAVSPITTSQTLFNLILGSMVAALAAQGAATAFVAGLAGAAGISV
GLGKVLVDILAGYGAGVAGALVAFKIMSGEPPTTDLVNLIPAILSPGALVGVVCAA
ILRHVGPEGAVOMNRLIAFASRGNHVSPTHVVPESDAAARVTAIILSLTVXLLR
RLHWVASECTPCSGSWLRDIDWICEVLSDFKTLWKAKLMPQLGPIPVSCORGR
GVWRGDGMHTRHCHGABEIGHVKNGTNRIVGPKTCRNMSGTFFPINAYTTGCTPLP
APNTFFALMRYSAEYVEIRRVGDHYVYTGMTADNLKPCQVPSPEFTELDGVLHR
FAPPKPLLRDEVSFRVLGHDYPVGSOLPCPEPDVAVLTSMLTDPDQHTAEAAARRL
ARGSPPEASQSOLAPSRLATCTTNHSDPAELTEANLLWRQEMGNITRVESEN
KVLIDFDPDLVAEDEREVSVAAILKSRRTFPALPIWARPDYNPPLLEKKKPDY
EPTVHGCPLPSPSPVPRKRTVVLSSTVSTALAEATKSPGSSSTSVTIGN
TASSEPPPSIICSDDAESCSMPLEGEFDPDLSGWSVTSVSSAGTEDVVCSSM
SYSTGALVTPCAAEQKLPINALNSLLRHNLVYITTSRACQOKVCKVDFRLQVL
DSHYQVLDKEVKAASVKANLLSVEBACSLTPHSKSKFGYKAKDVCHARKAVNL
INSVKLLLEASVTPIDTTWAKNEVFCVQPEKGCCKPARLIVPDLGVRFERKWL
DVNSKLLPVMSSGYGQSPQORVEFLVQAKSKTPMGFSYDTRCFDSTVTSBSDR
TEALYQCCDDJDPQAVAKSLERLYVGLPTNSNGENCGYRRCRAGSVLTSCGNT
LTCYKQAACRAARGLDCTWXXGDDLVWICESQVQDEAASLRATTEAMTRYSAFP
GDPPQYEDLELITSCSNSVAVHDGAKRYVYLTDRPTTDLARAWEATRHPTVNSW
LGNITMPATLWVRMLMTFFSVLIARDQEQALDEIYGCYISIEPLDLPPIIQLR
HGLSAFSLHSYSPGIEINRVAC"

ORIGIN

Alignment Scores:

Pred. No.: 8,05e-200 Length: 8791
Score: 3579.00 Matches: 671
Percent Similarity: 99.56% Conservative: 12
Best Local Similarity: 97.81% Mismatches: 3
Query Match: 98.92% Indels: 0
DB: 14 Gaps: 0

US-09-930-591-2 (1-686) x AY615798 (1-8791)

Qy	1	MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	3142	CTGGCGCCCATCGCGCGTATGCCARCAAGAGGGCCCTTTAGGATGCATAATCACC	3201
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla	40
Db	3202	AGTCTGACCGCGCGGACAAAACAGGTGGAGGTGAGATTGATGTCACCTGCT	3261
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThyHisGlyAla	60
Db	3262	GCCGACGATTTCTTGGCAACATGCATCAACGGGGTGTGCTGACTGTCTACACCGGGCC	3321
Qy	61	GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	3322	GGAAACAAGGACCATCGCGCTACCAAGGGTCTCTGTATCCAGATGTATTAATAGAC	3381
Qy	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	3382	CAAGACCTCGTAGGCTGGCCCGCTCCCAAGGCGCGCTCATTTGACGCGCTGCATTTGT	3441

Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg	120
Db	3442	GGCTCTCTCGGACCTTTACTCGTGCACGAGGCATGCTGATGTCAATTCCTCGCGCGGG	3501
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLeuGlySerSer	140
Db	3502	GGTGATAGAGGGGAGCGCTCTTTCGCCCCGGGCTATCTCTTACCTGAAAGGCTCTCTCG	3561
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	3562	GGGGGCCACCTGCTGTGCCCCGGGCGCACCGCTAGGCATATTTCAGGGCGCGGTGTGC	3621
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
Db	3622	ACCCGTGAGTGGCTAAAGCGGTGACTTTATTCCTGCTAGAGAACCTAGAGACACCATG	3681
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
Db	3682	AGGTCCCCGGGTGTTCACAGACAACCTCTCTCCCAACCGGGGGTGCCTCCAGAGCTTCCAGGTG	3741
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
Db	3742	GCCCACTGCTATCTCCCAACCGGCGAGTGGTAAGAGCACCAAGTCCCGCGCGGTATGCG	3801
Qy	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	3802	GCCCAAGGGCTACAAAGGTGCTGCTCAACCCCTCGGTGCGGTACATCGGGCTTCGCG	3861
Qy	241	AlaTyrMetSerLysAlaHisGlyLysAspProAsnIleArgThrGlyValArgThrIle	260
Db	3862	GCTTATATGTCGAAGGCCCGGGGTGATCCTTAACATCAGGACTGGGGTGAGGACAAT	3921
Qy	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
Db	3922	ACTACTGGCAGCCCATCATCGTACTCCACCTACCGCAAGTTCTTTCGCGCGCGGGTGC	3981
Qy	281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	3982	TCAGGGGTGCTTATGACATAATAATTAATTCGACGAGTGCCATTCACGAGTCAACATCC	4041
Qy	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	4042	ATCTTGGGCACTCGGCAACCGCTTTCGCAAGAGACTGCAGGGGCTAGACTGGTGTG	4101
Qy	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
Db	4102	CTGCGCACCGTACCCCTCGGGGTCTCATCTGTCCTCCCATCTTACATCGAGGAGGT	4161
Qy	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
Db	4162	GCTCTGTCCACCGGAGAGATCCCTTTTATGGCAAGGCTATCCCTTTGAGGCAATC	4221
Qy	361	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380
Db	4222	AAGGGGGGAGACATCTCATTTCTGTCATCTCAAGAAGAAATCGCAGACTCGCCGCA	4281
Qy	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	4282	AACTGTGCGGTGGGCATCAACCGCTGCTTACTACCGCGGCTTGAGCTGCTCCGTC	4341
Qy	401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
Db	4342	ATCCCGACCATGTCGACGCTGTCGTGCTAGCAACTGATGCTCTCATGACCGGCTTACC	4401
Qy	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
Db	4402	GGCAGCTTTCATCGGTGAGTAGACTGCAACACGCTGTGTCAACCCAGAGAGCTTCAGC	4461
Qy	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
Db	4462	CTTGACCTTCCTTACCATTTGAGACAACCACTTCCCAAGGATGTGTCTCCGCACT	4521
Qy	461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480

Db	4522	CAACCCCGGGGTAGACCCGGCAGGGGGAGCGCAGCATCTACAGATTGTGCGCGCGGGG	4581
Qy	481	GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	4582	GAGCGCCCTTCGGGCATGTTTGACTCGTCCGTCTCTGCGAGTGTGATGACCGGGTGT	4641
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrValArgLeuArgAlaTyrMetAsnThr	520
Db	4642	GCTTGGTATGAGCTCACACCCCGCAGACACAGTTAGCTTACGAGCATACATGAACAC	4701
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
Db	4702	CCGGGGCTTCCGTGTGCCAAGACCATTCTGAATTTTGGGAGGGCGTTTTTCACGGGTCT	4761
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	4762	ACCACATAGACGCCCACTTCTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTC	4821
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp	580
Db	4822	CTGGTGGCGTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCCCCCCCGTCTGGTGGAC	4881
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	4882	CAGATGTGGAAGTGTGATCCGCTCAAGCCACACTCCATGGGCCCAACACCTCTGCTA	4941
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	4942	TACAGGCTGGGTGGCGTTTCAGAATGAAGTCACTCCCTGACGCACCCAGTCAACAAGTACATC	5001
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	5002	ATGACATGCATGTCCGGCGACCTGGAGGTCTGACAGTAGTACCTGGGTGCTCGTGGTGGC	5061
Qy	641	ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	5062	GTTCTGGTGTCTTTGGCGCGCTACTGCCTATCCACAGGCTCGGTGTGTCATAATAGGCAGG	5121
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	5122	ATTGTCTTGCCGGAGCCGGCAGTCATCCCGCAGGGAGGTCTCTTACCGAGAGTTC	5181
Qy	681	AspGluMetGluGluCys	696
Db	5182	GATGAATGGAAGAGTGC	5199
RESULT 6			
LOCUS	I06434		
DEFINITION	Sequence 48 from Patent EP 0318216.	DNA	linear
ACCESSION	I06434		
VERSION	I06434.1	GI:590311	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 5360)		
AUTHORS	Houghton,M., Choo,Q.-L. and Kuo,G.		
TITLE	Nanbv diagnostics and vaccines		
JOURNAL	Patent: EP 0318216-A1 48 31-MAY-1989;		
FEATURES	Location/Qualifiers		
source	1..5360		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Alignment Scores:			
Pred. No.:	8,65e-200	Length:	5360
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0

```
Db 1950 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC 2009
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAala 380
Db 2010 AAGGGGGGAGACATCTCTCTGTCTATCAAGAGAGAGTGGAGCAACTCGCGCA 2069
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrArgGlyLeuAspValSerVal 400
Db 2070 AAGTGTGTGGATGGGATCAATCCGCTGCTACTACCGCGGTCTTGAGGTGTCGTC 2129
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2130 ATCCGACCGGGGAGTGTGTGTGTGTGGCAACCGATGCCCTCATACCGGCTATACC 2189
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2190 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTTCGATTCAGC 2249
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 2250 CTTGACCCCTACCTTCACCATTTGAGACAATCAGCTCCCGCAGGATGCTGTCTCCGCACT 2309
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2310 CAACGTGGGCGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGG 2369
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2370 GAGCGCCCTCCGCACTGTTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2429
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2430 GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTACGAGGCTACATGAACACC 2489
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 2490 CCGGGGCTTCGCGTGTGTCAGGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGGCTC 2549
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 2550 ACTCATATAGATGCCCATCTTCTATCCAGACAAGACAGTGGGAGAACCTTCTCTTAC 2609
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 2610 CTGCTAGCGTACCAAGCACCGTGTGCGCTAGGCTCAGGCCCTCCGCCATCGTGGAC 2669
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2670 CAGATGTGGAGTGTGTTGATTGCGCTCAAGGCCACCCCTCCATGGGCCAACACCCCTGCTA 2729
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 2730 TACAGATGGCGCGTGTTCAGAAATGAATACCTCTGACGACCCAGTCCACAAATACATC 2789
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 2790 ATGACATGCATGTGCGCGGACCTGGAGTGTCTACGAGGACCTGGTGTCTGTGGCGGC 2849
Qy 641 ValLeuAlaLeuAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 2850 GTCCTGGCTGCTTTGGCGCGTATTGCTGTCAACAGGCTGCTGGTTCATAGTGGCGAGG 2909
Qy 661 IleValLeuSerGlyLysProAlaIleProAspArgGluValLeuTyrArgGluPhe 680
Db 2910 GTGCTGTGTGCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 2969
Qy 681 AspGluMetGluGluCys 686
Db 2970 GATGAGATGGAGAGTGC 2987
```

RESULT 7
I09328

```
LOCUS I09328 5360 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 8 from Patent WO 8904663.
ACCESSION I09328
VERSION I09328.1 GI:587963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 5360)
AUTHORS Houghton, M., Choo, O.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 8 01-JUN-1989;
FEATURES
Location/Qualifiers
1..5360
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores: 8.65e-200 Length: 5360
Pred. No.: 3574.00 Matches: 672
Score: 3574.00
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0
US-09-930-591-2 (1-686) x I09328 (1-5360)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnLpThrArgGlyLeuGlyCysIleIleThr 20
Db 930 CTGGCGGCCCATCATCGGGGTACGCCCGCAGCAGACAAAGGGGCTCTTAGGGTGATATATCACC 989
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValCysTrpThrValThrAla 40
Db 990 AGCTTAATCGCGGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTTGTCTCACTGCT 1049
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThrHisGlyAla 60
Db 1050 GCCCAACCTTCTCTGGCAACGTCATCAATGGGTGTGTGACTGTCTACACGGGGCC 1109
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1110 GGACAGGAGGACCATCGCGTCCACCAAGGCTCTGTCTCATCAGATGTATACCAATGTATAGAC 1169
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1170 CAAGACCTTGTGGCTGTGGCGCTCCGCAAGGTAGCGCTCATTCACACCTTGACCTTGC 1229
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 1230 GGCTCTCGGACCTTTACCTGGTCACGAGGCAGCGCATGTCTCCGTGCGCGCGCG 1289
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 1290 GGTGATAGAGGGGACGCTGTGTGCGCGCGCGCCATTTCTACTTTGAAGGCTCTCTCG 1349
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAalaValCys 160
Db 1350 GGGGTTCCGTGTGTGTGTCGCCCGGGGACGCGCTGGGCATATTTAGGCGCGCGGTGTC 1409
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 1410 ACCGTGGAGTGGCTAAGCGGTGGACTTTTATCCCTGTGGAGAACCTTAGACACAACCATG 1469
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 1470 AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGTGCCCCCAGAGTCTCCAGGTG 1529
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAalaThrAla 220
Db 1530 GCTCACCTCCATGCTCTCCACAGGCGGCAAGACCAAGGCTCCCGGTGCTGCTATATGCA 1589
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
```

Db 1590 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 1649
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 1650 GCTTACATGTCGAAGGCTCATGGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAT 1709
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 1710 ACCACTGGCAGCCCACTACGTAATCCACCTACGCAAGTTCTTCCCGACGCGGGTGC 1769
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 1770 TCGGGGGCGCTTATGACATAATTTGTGACGAGTGCACCTCCACGATGCCACATCC 1829
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 1830 ATCTTGGGCATCGGCATGTCCTTGACCAAGCAGACTCGGGGGGAGACTGGTGTG 1889
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 1890 CTCCGACCGCCACCTCCCGGCTCCGTCACCTGTCGCCCATCCCAACATCGAGGAGTT 1949
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1950 GCTCTGTCCACCGGAGAGATCCCTTTTACGCAAGGCTATCCCTCCGAAGTAATC 2009
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 2010 AAGGGGGGAGACATCTCATCTTCTGTCTTCAAGAGAAAGTGCAGCACTCGCGCA 2069
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2070 AACTGTGTGTCATGGGCATCAATCGGTGCGCTACTACCGCGTCTTGACGTGTCGTC 2129
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2130 ATCCCGACCGCGGATGTGTGTCGTGCGCAACCGATGCCCTCATGACCGGCTATACC 2189
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2190 GCGCACTTCGACTCGGTGATAGACTGCATACGTGTGTACCCAGACAGTCGATTCAGC 2249
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 2250 CTTGACCTTACCTTACCATTGAGACAATCACTCCCGGAGTGTGTCTCCCGCACT 2309
QY 461 GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2310 CAACGTCGGGGCAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCAGGG 2369
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2370 GAGGGCCCTCCGGCATGTTGCACTCGTCTGCTGCTGAGTGTATGACGAGGCTGT 2429
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2430 GCTTGGTATGAGTCAAGCCCGCGAGACTACAGTTAGGTACGAGCGTACATGAACACC 2489
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 2490 CCGGGGCTCCCGTGTGCCAGGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCGTC 2549
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 2550 ACTCATATAGTCCCACTTCTATCCAGCAAGCAGAGTGGGAGAACCTTCTTTAC 2609
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 2610 CTGGTAGCGTACCAAGCCCGGTGCGCTAGGCTCAAGCCCTCCCTCCATCGTGGAC 2669
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2670 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCCAACCTCCATCGGCGCAACACCCCTGCTA 2729

QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 2730 TACAGACTGGCGCTGTTCAGAATGAATCACTCAGCAGCCAGTCAACCAATACATC 2789
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 2790 ATGACATGATGTGCGCCGACCTGGAGGTGCTCACGAGCACCTGGGTGCTGTGGCGC 2849
QY 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 2850 GTCCTGGCTGCTTGGCCGCGATTCCTGTCAACAGGCTGCGTGTGTCATAGTGGCAGG 2909
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 2910 GTCGTCTTGTCCGGGAAGCGGCAATCATCATCTGACAGGGAAGTCTCTTACCAGAGTTC 2969
QY 681 AspGluMetGluGluCys 686
Db 2970 GATGAGATGGAAGAGTGC 2987
RESULT 8
106440
LOCUS 106440
DEFINITION Sequence 54 from Patent EP 0318216.
ACCESSION 106440
VERSION 106440.1 GI:590312
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 6785)
AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.
TITLE Nanbv diagnostics and vaccines
JOURNAL Patent: EP 0318216-A1 54 31-MAY-1989;
FEATURES Location/Qualifiers
source 1..6785
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.15e-199 Length: 6785
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0
US-09-930-591-2 (1-686) x 106440 (1-6785)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1203 CTGGCGCCCATCATCGGCGTACGCCAGCAGACAGGGGCTCTTAGGTGTCATATACCC 1262
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 1263 AGCTTAATGGCGCGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCACTGCT 1322
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 1323 GCCCAAACTTCTTCGCAACGTCATCAATCAATGGGGTGTGTGGACTGTCTACACGCGGCC 1382
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1383 GGAACGAGGACCATCGCGTCACCAAGGTCCTGTATCCAGATGATATACCAATGTAGAC 1442
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1443 CAAGACTTGTGGCTGGCCCGCTCCGCAAGGTAGCGCTCATTTGACACCTGCACTGC 1502
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120

Db 1503 GGCTCCTCGGACCTTTACCTGGTCACGAGGACGCCGATGTCATTCCTCGTGGCGGGCGG 1562
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLeuGlySerSer 140
Db 1563 GGTGATAGCAGGGCAGCGCTGCTGCGCCCGGCCCATTTCTACTCTGAAAGGCTCCTCG 1622
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 1623 GGGGGTCCGCTGTTGTGCGCCCGGGGACGCCGCTCGGCATATTTAGGGCGCGGTGTGC 1682
Qy 161 ThrArgGlyValAlaLeuAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 1683 ACCCGTGGAGTGGCTAAGCGGTGACCTTTATCTCTGGGAGAACCTTAGACAAACCATG 1742
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 1743 AGGTCCCGGTTGTTACCGGATAACTCTCTCCACCAGTAGTGCCCGCAGAGCTTCCAGGTG 1802
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 1803 GCTCACTCTCCATGCTCCACAGGAGCGGCAAAAAGCACCAAGGTCGCCGCTGCATATGCA 1862
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 1863 GCTCAGGCTATAAGGCTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGCTTTGGT 1922
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 1923 GCTTACATGTCAGGCTCATGGATCGATCTTAACATCAGACCGGGGTGAGAACAAATT 1982
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 1983 ACCACTGGCAGCCCATCACCTACTCCACCTACGCGCAAGTTCTTGGCGAGCGGGTGC 2042
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2043 TCGGGGGCGGCTTATGACATAATAATTGTGACAGTGCCTCCACGGATGCCACATCC 2102
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2103 ATCTTGGGCATCGCAGCTGCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTG 2162
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2163 CTCGCCACCGCCACCCCTCGGGTCCGTCCTACTGTGCCCATCCCAACATCGAGAGGTT 2222
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2223 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAAGTAACT 2282
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2283 AAGGGGGGAGACATCTCTCTGTCTTCAAGAGAGAGTCCGAGAACTCCCGCA 2342
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2343 AAGCTGTGTCATTGGGCATCAATCGCGTCTACTACGCGGTCTTTGAGCTGTCGTC 2402
Qy 401 IleProThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2403 ATCCCGACCGCGCGATGTTGCTGTCGTCGCAACCGATGCCCTCATGACCGGTATACC 2462
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2463 GGCAGCTTCGACTCGGTGATAGATGCAATACGTGTGTACCCAGACAGTCGATTTCCAGC 2522
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 2523 CTTGACCTTACCTTACCATTTGACACATCAGCTCCCGCAGGATGCTGTCTCCGCACT 2582
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2583 CAACGTGGGGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGG 2642

Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2643 GAGCGCCCTCCGCATGTTCCGACTCGTCCGTCTCTGTGAGTGTATGATGCGAGGCTGT 2702
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2703 GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 2762
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 2763 CCGGGGCTTCCGTTGTCCAGGACCACTTGAATTTTGGAGGGCGTCTTTACAGGCTC 2822
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 2823 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 2882
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrPhe 580
Db 2883 CTGGTAGCTACCAAGCCACCGTGTGGCTCAGGCTCAAGCCCTCCCCCATCGTGGAC 2942
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2943 CAGATGTGAAGTGTGTGATTCGCCTCAAGCCCACTTCCATGGGCCAACCCCTGCTA 3002
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3003 TACAGACTGGGCGCTGTTCCAGATGAATCACCCTGACGACCCAGTCACCAATACATC 3062
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 3063 ATGACATGCTATGTCGGCGCACCTGAGGTGCTCAGGACACCTGGTGTCTGTTGGCGC 3122
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValIleArg 660
Db 3123 GTCTGTGCTGTTTGGCGCGGTATGTCTGTCAACAGGCTGCGGTGCATAGTGGGCGG 3182
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3183 GTCGTCTGTCCGGAGCGGCAATCATACCTGACAGGGAGTCTCTTACCGAGATTC 3242
Qy 681 AspGluMetGluGluCys 686
Db 3243 GATGAGATGAAGAGTGC 3260

RESULT 9
I09329
LOCUS I09329 6785 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 10 from Patent WO 8904669.
ACCESSION I09329
VERSION I09329.1 GI:587964
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 6785)
AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 10 01-JUN-1989;
FEATURES Location/Qualifiers
source 1. 6785
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.15e-199 Length: 6785
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x I09329 (1-6785)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db : : : : :
1203 CTGGCGCCCATCACGGGTACGCCAGCAGACAAAGGGGCTCTCGGTGCATATCACC 1262

QY 21 SerLeuThrGlyArgAspLysAenGlnValGluGlyGluValGlnIleValSerThrAla 40
Db : : : : :
1263 AGCCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCT 1322

QY 41 AlaGlnThrPheLeuAlaThrCysIleAenGlyValCysTrpThrValTyrHisGlyAla 60
Db : : : : :
1323 GCCCAACCTTCTCGCAACAGTCATCAATGGGGGTGTCTGGAGCTGTCTACCAAGGGGCC 1382

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAenValAsp 80
Db : : : : :
1383 GGNACGAGGACCATCGCGTCACCAAGGGTCTGTCTCATCCAGATGTATACCAATGTAGAC 1442

QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db : : : : :
1443 CAAGACCTTGTGGCTGGCCGCTCCGCAAGGTAGCGGCTCAATTGACACCTCGACTTGC 1502

QY 101 GlySerSerAspLeuTyrIleValThrArgHisAlaAspValIleProValArgArgArg 120
Db : : : : :
1503 GGCTCTCGGACCTTACCTGGTTCAGAGCACGCCGATGTCAATCCCGTGGCCGGCGG 1562

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db : : : : :
1563 GGTGATAGACGGGGACGCTGTCTGCCCGGCGCCCATTTCTTCTTGAAGGCTCTCTCG 1622

QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db : : : : :
1623 GGGGGTCCGCTGTGTGGCCCGCGGGGACCGCTGGGCATATTTAGGGCCGGGGTGTGC 1682

QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db : : : : :
1683 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGACACAACCATG 1742

QY 181 ArgSerProValPheSerAspAenSerSerProProAlaValProGlnSerTyrGlnVal 200
Db : : : : :
1743 AGGTCCCCGGTGTTCAGGNATACTCTCTCCACAGTAGTGCCCGACAGCTTCCAGGTG 1802

QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db : : : : :
1803 GCTCACTCCATGCTCCACAGGCAGCGCAAGAACCAAGCTCCCGGTGCATATGCA 1862

QY 221 AlaGlnGlyTyrLysValLeuValLeuAenProSerValAlaAlaThrMetGlyPheGly 240
Db : : : : :
1863 GCTCAGGGCTATTAAGGTGTCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGGCTTTGGT 1922

QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db : : : : :
1923 GCTTACATGTCGAAGGCTCATGGATCGATCCTTAACATCAGGACCGGGGTGAGAACATTT 1982

QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db : : : : :
1983 ACCACTGGGAGCCCATCAGTACTCCACCTACGGCAAGTTCCTTGGCCGAGCGGGTGC 2042

QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db : : : : :
2043 TCGGGGGCGGCTTATGACATAATAATTGTGAGGAGTGGCCACTCCACGGATGCCACATCC 2102

QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db : : : : :
2103 ATCTTGGGCATCGGCACTGCTTGTACCAAGCAGAGACTCGGGGGGCGAGCTGGTTGTG 2162

QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAenIleGluGluVal 340
Db : : : : :
2163 CTCGCCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCATCCCAACATCGAGGAGGTT 2222

QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db : : : : :
2223 GCTCTGTGCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 2282

RESULT 10
LOCUS AR118696
DEFINITION Sequence 74 from patent US 6150087.
ACCESSION AR118696

AR118696 7310 bp DNA linear PAT 16-MAY-2001
Sequence 74 from patent US 6150087.
AR118696

QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db : : : : :
2283 AAGGGGGGCGAGACATCTCATCTCTCTCTCAATCAAGAGAGAGTGCAGCACTCGCGCA 2342

QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db : : : : :
2343 AAGCTGGTGCATTTGGGCATCAATCCGTGGCTACTACCGGGTCTTACGCTGTCGTC 2402

QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db : : : : :
2403 ATCCCGACGACGGCGATGTTGTCGTGGCAACCGATGCCCTCATGCCGCTATACC 2462

QY 421 GlyAspPheAspSerValIleAspCysAenThrCysValThrGlnThrValAspPheSer 440
Db : : : : :
2463 GGCAGCTTCCACTCGGTGATAGACTGCAATACGTGTGTCCACGACAGTCGATTTCCAGC 2522

QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db : : : : :
2523 CTGTACCCCTTACCTTACCAATTGAGACATACGCTCCCCAGGATGCTGTCTCCCGACT 2582

QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db : : : : :
2583 CAACGTGCGGGCAGGACTGCGAGGGGGAAGCCAGGCATCTACAGATTGTGGCACCGGG 2642

QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db : : : : :
2643 GAGCGCCCTCGGCAATGTTCCGACTCGTCCGTCTCTGTGAGTGTATGACGAGGCTGT 2702

QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAenThr 520
Db : : : : :
2703 GCTTGTGTATGAGCTACGCCCGCGGAGACTACAGTTAGGCTACGAGCGTACATGAACAC 2762

QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db : : : : :
2763 CCGGGGCTTCCCGTGTGCCAGACCATCTTGAATTTGGAGGGCGCTTTTACAGGCTC 2822

QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAenLeuProTyr 560
Db : : : : :
2823 ACTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 2882

QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db : : : : :
2883 CTGGTAGCGTACCAAGCCCGGTGCGCTAGGGCTCAAGCCCTCCCCATCGTGGGAC 2942

QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db : : : : :
2943 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCCAACCTCCATGGGCCAACACCCCTGCTA 3002

QY 601 TyrArgLeuGlyAlaValGlnAenGluValThrLeuThrHisProValThrLysTyrIle 620
Db : : : : :
3003 TACAGACTGGCGCTGTTCAAGATGAATCACCTGACGCCACCCAGTCACCAATACATC 3062

QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db : : : : :
3063 ATGACATGATGTGCGCCGACCTGGAGGTGCTCAGCAGACACCTGGGTGCTGTGGGGC 3122

QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db : : : : :
3123 GTCTGGCTGCTTGGCCGGGTATTCGCTGTCAACAGGCTGCGTGGTCATAGTGGGCGAG 3182

QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db : : : : :
3183 GTGCTTGTGTCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTCTACCGAGAGTTC 3242

QY 681 AspGluMetGluCys 686
Db : : : : :
3243 GATGAGATGAAGAGTGC 3260


```
Db 3528 TACAGACTGGCGCTGTTAGAAATCAATCACTCAGCGACCCAGTCACCAATACATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3588 ATGACATGTCATGTCGGCGGACCTGGAGGTGTCACGAGCACCTGGGGTCTCGTTGGGGGC 3647
Qy 641 ValLeuAlaLeuAlaAlaValCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTGTGCTGTTGGCGCGGTATGCTGTCACAGCGTCGGTGGTTCATAGTGGCAGG 3707
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuValArgGluPhe 680
Db 3708 GTGCTGTTGTCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 3767
Qy 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785

RESULT 11
LOCUS I09331 7310 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 15 from Patent WO 8904669.
ACCESSION I09331
VERSION I09331.1 GI:587966
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 7310)
AUTHORS Houghton,M., Choo,Q.-K. and Kuo,G.
JOURNAL Patent: WO 8904669-A 15 01-JUN-1989;
FEATURES
Location/Qualifiers
1..7310
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 1..26e-199 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x I09331 (1-7310)
Qy 1 MetAlaProIleThrAlaValAlaGlnInThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 CTGGCGCCCATCAGCGGTACGCCAGCACAAAGGGCCCTCTAGGGTCATATACCC 1787
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 1788 AGCCTAACTGCGCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTCAACTGCT 1847
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThrHisGlyAla 60
Db 1848 GCCCAACCTCTCTGGCAACGTGCATCAATGGGGTGTCTGGAGTGTCTACACAGGGGCC 1907
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1908 GGAACGAGGACCACCGCTCACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 1967
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1968 CAAGACCTTGTGGCTGGCCGCTCCGCAAGTAGCGCGCTCAATGACACCCCTGCATTGC 2027
Qy 101 GlySerSerAspLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 2028 GGCTCTCGGACCTTTACCTGGTTCACGAGCAGCCGATGTCAATCCCGTCCCGCGCGG 2087
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
```

```
Db 2088 GGTGATAGCAGCGGAGCTGTGTGCGCCCGCCCATTTCTTCTTGAAGGCTCCTCG 2147
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 2148 GGGGGTCCGCTGTGTGTCGCCCGGGGACGCGCGTGGGCATATTTAGGGCGGGGTGTC 2207
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 2208 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGAACCATG 2267
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGTCCCCCGGTGTTACGGGATACTCTCTCTCCACCATAGTAGTCCCGAGCTTCCAGGTG 2327
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIleValAla 220
Db 2328 GCTCACCTCCATGCTCCACAGCAGCGCAAGCAACCAAGGTCCCGGTGCATATGCA 2387
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGCTATAGGTGTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 2447
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 2507
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCATGGCAGCCCATCATCGTACTCCACTAGCGCAAGTTCCTTCCGACGCGGGTGC 2567
Qy 281 SerGlyGlyAlaValAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGGCGCTTATGACATATATATTTGTGACAGTGCACCTCCAGGATGCCACATCC 2627
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTTGGCATCGGCATGTCCTTGACCAAGCAGAGACTCGCGGGCGAGACTGGTTGTG 2687
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTCGCCACCGCCACCCCTCGGGCTCCGTCACCTGTCGCCCATCCCAACATCGAGGAGTT 2747
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2748 GCTCTGTCCACCACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTTCGAAGTAATC 2807
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCTCATCTTCTGCTCAATCAAGAAGAGTGCAGCAACTCGCGCA 2867
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGTGTCGATTTGGGCTCAATGCCGTGGCTACTACCGGGTCTTGACGTTCCGTC 2927
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCCGACGAGCGCGATGTTGTCGTGGGCAACCGATGCCCTCATGCCGCTATACC 2987
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GGCAGCTTCGACTCGGTGATAGACTGCAATACGTGTGTCAACGAGAGTCCGATTTTCAGC 3047
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTGACCCCTACCTTCACTTACCAATGAGCAATCAGCTCCCGGATGCTGTCTCCGCACT 3107
Qy 461 GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 3108 CAACTGCGGGCAGGACTGGCAGGGGAGCCAGGCACTTACAGATTTGTGGCACCGGGG 3167
Qy 481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500
```

Db	3168	GAGCGCCCTCGGCATGTTTCAGACTGCTCGCTCGCTCTGAGTGTATGACGCGAGCTGT	3227
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	3228	GCTTGGTATGAGCTCAGCGCCGCGAGACTACAGTTAGCTACGAGCGTACATGAACACC	3287
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu	540
Db	3288	CCGGGGCTTCCTCGTGTGCAGGACCATCTTGAATTTTGGGAGGGCGCTTTTACAGGCTC	3347
Qy	541	ThrHisTleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	3348	ACTCATATAGATGCCACTTTCTATCCAGAACAGCAGTAGTGGGAGAACCTTCTCTTAC	3407
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp	580
Db	3408	CTGGTAGCTACCAAGCCACCGTGTGCGCTAGGAGCTCAAGCCCTCCCTCCCATCGGGAC	3467
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	3468	CAGATGTGGAAGTGTTCGCTCGCTCAAGCCACCTCCATGCGGCCAACACCCCTGCTA	3527
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	3528	TACAGATGGGGCGCTGTTCAGAAATCAACCTCGACGACCCAGTCACCAATACATC	3587
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	3588	ATGACATGATGTGGCGGACCTGGAGCTGTCACGAGCACCTGGGTCTGTTGGCGGC	3647
Qy	641	ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	3648	GTCCTGGCTGTTTGGCGCGCTATTGCTGTCAACAGGCTCGTGGTTCATAGTGGCAGG	3707
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	3708	GTCCTGTGTCGGGAGCGGCAATACATCTCACAGGGAAGTCTCTACCGAGAGTTC	3767
Qy	681	AspGluMetGluGluCys 686	
Db	3768	GATGAGATGGAGAGTGC 3785	
RESULT 12			
HPCPOLYP			
LOCUS	HPCPOLYP 7310 bp ss-RNA linear VRL 02-AUG-1993		
DEFINITION	Hepatitis C virus polyprotein gene, partial cds.		
ACCESSION	M32084		
VERSION	M32084.1 GI:329875		
KEYWORDS	polyprotein.		
SOURCE	Hepatitis C virus		
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1 (bases 1 to 7310)		
AUTHORS	Choo, Q.-L., Richman, K. and Han, J.		
TITLE	The nucleotide sequence of the Hepatitis C viral genome		
JOURNAL	Unpublished (1990)		
COMMENT	Original source text: Hepatitis C virus, cDNA to viral RNA, clones K9-1 through 15e, isolated from chimpanzee (individual 910) blood plasma. Draft entry and printed sequence for [1] kindly submitted by M.Houghton, 22-FEB-1990. Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608.		
FEATURES	Location/Qualifiers		
source	1..7310		
	/organism="Hepatitis C virus"		
	/mol_type="genomic RNA"		
	/db_xref="taxon:11103"		
	<1..7310		
	/note="polyprotein"		
	/codon_start=3		
CDS	/protein_id="AAA45677.1"		
	/db_xref="GI:329876"		

/translation="GCBERLASCRPLTDFDQGWGPISYANGSGDPQDRPYCWHYPPKPC
TWNSTGFTKVCYCAPPCVIGAGNNTLHCPTDCFRKHPDATYSCRGSPWITPRCLVD
YPRVLMHPCTINVTIPIKIMYGVGVHREAAQWTRBSCRDCSLASWAIKWEYVLLPLLL
TQOMVLPCTFTTLPALSTGLJHLQHNVDVQVLYGVGSSISWAIKWEYVLLPLLL
ADARVCSCLMMLLSIQAEAALENLVLINAAASLAGTHGLVFLVFFCFANVLYKKNVP
GAVTYFGMPLLLLLLALPQRAYLDEVAASCGVVLVGLMALTITSPYKRYISWC
LWMLYFLTRVEAQLHVMIPPLNVRGGRDAVILLMCAVHTPLVFDITKLLAVGLPLW
ILQASLLKVPYFVRVQGLLRICALAKMIGHYVQMWIKLGLATGTIVYVNHLLTPRD
WAHNLRLDAVAVPEVPVSOMETKLIITWADTAACGDIINGLPSVARRRGRLGLPAD
GVMSKWRLLAPITAYAOQTGRLGCLITSLTGRDKNOVEGEVQIVSTAQTFLATCI
NGVCWTYHGAGTRTIAAPKGPVQMTINVDQDLVMPAPQSGSLTPTCTCGSSLDL
VTRHADVPIVRRREGDSGLSPRPISYILKSGSGPLCPAGHAGVLPFRAAVCTRGVA
KAVDIFIVENLETTRMSPVFNSSPPVQSFQVAHLHAPTSGKSTKVAAPAAQOG
YKVLNVPSVAATLFGAYMSKAGHIDPNTIGVTTITGSPITVSTYKGLDGGCS
GGADVITICDCHSTDATSLIGCTVLDQAEATAGARLVLATATPPGSVTPVPHNIE
VALSTGEIPYGCATPLEVIGKGRHLIFCHSKKDELAALVALGINAVAYRGLD
VSVITSGDVVVVAITDALTGYTDFDSVIDCNCTVCTVDFTITITILPQD
AVSRTQRTGRGKPGIYRFVAFGRPSGMFDSVLCCEYDAGCAGMYELTPAETTVR
LRYMNTPGLPQODHLEFWEGVFTGLTHDAHLSQTKQSGENLPVLAQYATCAR
AQAAPPDMDQWKCLIRLKPTLHGPTPLLYRLGAVONEITLTHKPVKYIMTCSADLE
VVTSTVVLVGVLAALAAAYCLSTGCVVIIGRVVLGKPAITPDREVLRYRDEMBECS
QHLPYIQGMMLEAQKQKALGLLOTSARQAEVIAVOTNWOKLETFWAKHMFIS
GIQYLAGSLTPGNPALASMAFTAAVTSPLTSTQTLNLFILGGWVAQAALAPGAATA
FVQGLAGAAIGSVGLKVLITDILAGYAGVAGALVAFKIMSGVSPSTEDLVNLLPAI
LSPCALVGVVCAAILRRHVGPGEQVQMMNRLIAFASRGNHVSPTHYVPESDAAARV
TAILSSVITQLLRLHWTISSECTTPCSGSLRDINDWICEVLSDFKWLKALMPQ
LPGIPFVSCORGKGVWVVDGIMHTRCHCGAEITGHVKNGTMRIVGPRTCNMWSGTF
PINATYGPCTPLPAPNYTALWRVSAEEVEIRQVGDHYVTVGTNTDNLKCPQVPS
PERFTELDGVLRLHFAAPPKTLREEVSFRVGLHVEPVGSQLPCEPDPVVALTSMLT
QKQVTRKARLVTHINSWKVLEKAAASKVLANLLSVEACSLTPPHSKAKFGYV
ADRVCHARKAVTHINSWKVLEKAAASKVLANLLSVEACSLTPPHSKAKFGYV
PDGLRVYCEKMAVLDVVKLPVAVMGSSYGFQGVQGVQVFLVQAMKSKTTPMGFYV
TRCFDSTVTESDIRTEBAIYQCDOLDPOARVAIKSLTERLVVGGPLNTRGECGYRR
CRASGLVITSCGNTLTCYIKARACRAAGLQDCTMLVCGDDLVVICBSAGVQEDASL
RAPTEAMTRYSAPPDPPQPEYDELITSCSSNVSVVAHDGKGVKIVITRDPPTPLAR
AAWETARHTFVSNLGNIMFAPTLWARMILMTHFFSVLIARDQLEQALDECIYGCY
SIEPLDLPPIQL"

ORIGIN

Alignment Scores:			
Pred. No.:	1.26e-199	Length:	7310
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	14	Gaps:	0
US-09-930-591-2 (1-686) x HPCPOLYP (1-7310)			
Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	1728	CTGGCGCCCATCAGCGGTACGCCACAGACAGGGGCGCTCTTAGGTGCATATACACC	1787
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla	40
Db	1788	AGCCTAACTGCGCGGACAAAAACCAAGTGAGGGTCCAGATGTGTCTCAACTGCT	1847
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThiGlyAla	60
Db	1848	GCCCAAACTCTCTGGCAACGTGCATCAATGGGGTGTGCTGGAGCTGTCTACACGGGGCC	1907
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	1908	GGAACGAGGACCATCGGCTCACCACAGGGTCTCTGTCATCCAGATGTATACCAATGTAGAC	1967
Qy	81	GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100

```
Db 1968 CAAGACCTTGTGGGCTGGCCGCTCCCGAAGGTAGCGCTCAATTGACACCCCTGCACCTTGC 2027
QY 101 GlySerSerAspLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 2028 GGTCTCTCGGACCTTTACCTGGTTCACAGGACGCGGATGTCTATTCCTCGCGCGCGG 2087
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 2088 GGTGATAGCAGGGGACCTGCTGTGCGCCCGGCCCATTTCTTACTTGAAGGCTCTCTCG 2147
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaIaValCys 160
Db 2148 GGGGGTCCGCTGTGTGTCGCCCGGGGACGCGGTCGGGCATATTTAGGGCGCGGTGTGC 2207
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 2208 ACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTTAGACAGAACCATG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGTCCCGGTGTTCAGGATACTCTCTCCACCATGAGTGTGCCCGCAGAGCTTCCAGGTG 2327
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIaTyrAla 220
Db 2328 GCTCACCTCCATGCTCCACAGGCGAGCGGCAAGACCAAGGTCGCGGCTGCATATGCA 2387
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaIaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGGCTTTGGT 2447
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCGAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAT 2507
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCACTGGCAGCCCATACGTACTCCACCTACGCAAGTTCCTTGGCGCAGCGGGTGC 2567
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGGGCGCTATGACATATATATTTGTGACAGTGCACCTCCACGGATGCCATCC 2627
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTTGGGCATCGCATGCTCTTACCAGCAGAGACTCGCGGGGCGAGACTGGTTGTG 2687
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGAGGTT 2747
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2748 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAGGCTATCCCTCTCGAAGTAATC 2807
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaIa 380
Db 2808 AAGGGGGGAGACATCTCTCTCTGTCTTCAATCAAGAAAGTGCAGCAACTCGCGCA 2867
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGTGTCATTTGGGCATCAATCGGTGGCTTACTACCGCGGTCTTGACGTGTCGTC 2927
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCCGACCGCGGATGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGCTATACC 2987
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GGGCACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCAGACAGTCCGATTTTCAGC 3047
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCCCTACCTTCACATTGAGACAAATCACGCTCCCCAGGATGCTGTCTCCCGACT 3107
```

```
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyLysProGlyLysPheValAlaProGly 480
Db 3108 CAACGTGGGGAGGACTGGCAGGGGAGAGCCAGGCACTTACAGATTTGTGGCACCAGGG 3167
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCGGCATGTTGACTCGTCCGTCCTCTGTGAGTGTCTATGACGAGGCTGT 3227
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 3228 GCTTGTGTATGACTCAGCGCCGCGAGACTTACAGTTAGGCTAGAGCTACATGAACACC 3287
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCCGGTGGCCAGGACCATCTTGAATTTTGGAGGGCGCTTTTACAGGCGCTC 3347
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAGAACTTCTCTTAC 3407
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 3408 CTGCTAGCTTACCAAGCCACCGGTGTGCTAGGGCTCAAGCCCTCCCTCCATCGTGGAC 3467
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGAAGTGTGTTGATTTGCCTCAAGCCACCCCTCATGGGCGAACACCCCTGCTA 3527
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrTyrIle 620
Db 3528 TACAGACTGGGCGCTGTTTCAAGATGAATCACCCTGACGACCCAGTCAACAAATACATC 3587
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGly 640
Db 3588 ATGACATGATGTGCGCCGACCTGGAGGTGCTGACAGACACCTGGGTGCTGTGGCGGC 3647
QY 641 ValLeuAlaAlaLeuAlaIaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTCTGGCTGCTTTGGCCCGTATTGCTGTCAACAGGCTGGGTGCTATAGTGGGAGG 3707
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTCTGCTTGTCCGGGAGGCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 3767
QY 681 AspGluMetGluGluCys 686
Db 3768 GATGATGGAAGAGTGC 3785

RESULT 13
LOCUS I08294 9185 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent EP 0388232.
ACCESSION I08294
VERSION I08294.1 GI:588994
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 9185)
AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.
TITLE NANBV diagnostics and vaccines
JOURNAL Patent: EP 0388232-A1 1 19-SEP-1990;
FEATURES Location/Qualifiers
source 1..9185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores: 1,67e-199 Length: 9185
Pred. No.: 3574.00 Matches: 672
Score: 99.42% Conservative: 10
Percent Similarity:
```

Best Local Similarity: 97.96%		Mismatches: 4	
Query Match: 98.78%		Indels: 0	
DB: 6		Gaps: 0	
US-09-930-591-2 (1-686) x I08294 (1-9185)			
Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	3395	CTGGCGCCCAATCACGGCGTACGCCAGCAGACAAGGGCGCTCTTAGGGTGCAATATCAC	3454
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla	40
Db	3455	AGCCTAACTGCGCGGACAAAACAAGTAGAGGGTGAGGTCCAGATGTGTCAACTGCT	3514
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla	60
Db	3515	GCCCAAACTTCCTGGCAACGTGCATCAATGGGGTGTGTGACTGTCTACCAACGGGGCC	3574
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	3575	GGAAACGAGGACCATCGCGTCAACCAAGGGTCTCTGTCTATCCAGATGTATACCAATGTAGAC	3634
Qy	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	3635	CAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTTGACACCCCTGCACATTGC	3694
Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg	120
Db	3695	GGCTCTCGGACCTTACCTGGTCACGAGGACACGCCGATGTCTATCCCGTGGCGGGGG	3754
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer	140
Db	3755	GGTGATAGCAGGGCAGCTGCTGTGCGCCCGGCCCATTTCTACTTTGAAAGGCTCCTCG	3814
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	3815	GGGGGTCCGCTGTGTGCCCCCGGGGACGCCGTGGGCATATTTAGGGCGCGGTGTC	3874
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
Db	3875	ACCGTGGAGTGGCTAAGCGGTGGACTTTATCTCTGTGGAGAACCTTAGAGACAACCATG	3934
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
Db	3935	AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGCCGCCAGAGCTTCCAGGTG	3994
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
Db	3995	GCTCACCTCCATGCTCCACAGGACGCGCAAAAGCACCAAGGTCCCGGCTGCATATGCA	4054
Qy	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	4055	GCTCAGGGCTATAGGTGTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGCTTTGGT	4114
Qy	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	4115	GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGACCGGGGTGAGAACAAAT	4174
Qy	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys	280
Db	4175	ACCACCTGGCAGCCCATCAGCTACTCCACCTACGGCAAGTTCTTTGCGCAGCGGGGTGC	4234
Qy	281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	4235	TCGGGGGCGCTTATGACATAAATAATTTGTGACAGGTGCCACTCCAGGATGCCACATCC	4294
Qy	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	4295	ATCTTGGGCATCGGACTGTCTTGTACCAACGACGAGACTGGCGGGGCGAGACTGGTTGTG	4354
Qy	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
Db	4355	CTGGCCACCGCACCCCTCCGGGCTCCGTCACTGTGCCCCCATCCCAATCCGAGGAGGTT	4414
Qy	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
Db	4415	GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGTATCCCTCGAAGTAATC	4474
Qy	361	LysGlyGlyArgHisIleLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380
Db	4475	AAGGGGGGAGACATCTCATCTCTGTGTCATTCAAAAGAAAGTCGAGAACTCCGCCGA	4534
Qy	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	4535	AAGTGTGTGCATTTGGGCATCAATGCGTGGCTACTACCGGGGTCTTGACGTGTCCGTC	4594
Qy	401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
Db	4595	ATCCCGACACGCGCATGTTGCTGCTGGCAACCATGCTCATGACCGGTATATACC	4654
Qy	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
Db	4655	GGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCTACCCAGACAGTCGATTTCCAGC	4714
Qy	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
Db	4715	CTTGACCTTACCTTCCACTTGGACAAATCAGCTCCCCAGGATGTGTCTCCCGCACT	4774
Qy	461	GlnArgArgGlyArgThrGlyArgGlyPheProGlyIleTyrArgPheValAlaProGly	480
Db	4775	CAACGTGGGGCAGGACTGCGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGGG	4834
Qy	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	4835	GAGCGCCCTCCGCGATGTTGCACTCGCTCGCTCTCTGTGAGTGTATGACGACGAGTGT	4894
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	4895	GCTTGGTATGAGCTCAGCCCCCGGAGACTACAGTTAGGCTACGAGCGTACATGAACACC	4954
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
Db	4955	CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAAATTTTGGGAGGGCGCTTTTACAGGCCTC	5014
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	5015	ACTCATATAGATGCCACTTTCTATCCACAGACAAAGCAGAGTGGGGAGAACCTTCTCTTAC	5074
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp	580
Db	5075	CTGTAGCGTACCAGCACCGGTGTGGCTAGGGCTCAAGCCCCCTCCCCCATCTGGGGAC	5134
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	5135	CAGATGTGGAAGTGTGTGATTGCTTCAAGCCCCACCTCCATGGGCCAACACCCCTGCTA	5194
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	5195	TACAGACTGGGGCTGTTTCAAGAAATGAAATCACCTGACGCAACCCAGTCACCAATAATC	5254
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	5255	ATGACATGCAATGTGGCGGACCTGGAGGTGCTCAGAGCACCTGGGTGCTGTGGCGGCG	5314
Qy	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	5315	GTCTGTGCTCTTGGCGCGGTATTGCTGTCAACAGGCTGCGTGGTCATAGTGGGCGAGG	5374
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	5375	GTGCTCTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC	5434
Qy	681	AspGluMetGluGluCys	686
Db	5435	GATGAGATGGAAGAGTGTC	5452

```
RESULT 14
LOCUS BD091382 9185 bp DNA linear PAT 27-AUG-2002
DEFINITION HCV cultivation method in eucaryotic cells.
ACCESSION BD091382
VERSION BD091382.1 GI:22636993
KEYWORDS JP 2001314192-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 9185)
AUTHORS Weiner,A.J., Steimer,K.S. and Houghton,M.
TITLE HCV cultivation method in eucaryotic cells
JOURNAL Patent: JP 2001314192-A '3 13-NOV-2001;
CHIRON CORP
OS Homo sapiens (human)
PN JP 2001314192-A/3
PD 13-NOV-2001
PF 15-MAR-2001 JP 2001075114
PR 25-AUG-1989 US 398667
PI AMY J WEINER, KATHELYN S STEIMER, MICHAEL HOUGHTON PC
CI 2N15/09, CI 2N5/10, CI 2N7/00// (CI 2N7/00, CI 2R1/93), CI 2N15/00, PC
CI 2N5/00
CC HCV cultivation method in eucaryotic cells
FH Key Location/Qualifiers
FT source 1. 9185
FT /organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1. 9185
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.67e-199 Length: 9185
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0
US-09-930-591-2 (1-686) x BD091382 (1-9185)
QY 1 MetAlaProfileThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
3395 CTGGCGCCCATACGGGTACGCCCGTACGACCAAGGGGGCTCTCTAGGGTGCATAATCACC 3454
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGluValGlnIleValSerThrAla 40
Db AGCTTAATGCGCGGACAAACCAAGTGGAGGGTGAGTCCAGATTGTGTCACTGCT 3514
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db GCCCAAAACCTTCTCTGGCAAGTGCATCAATGGGGTGTCTGGACTGTCTACCAAGGGGCC 3574
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db GGAAACGAGGACCATCGCGTCACCAAGGGTCTGTATCCAGATGATATACCAATGAGAC 3634
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db CAGACCTTGTGGCTGGCCCGCTCCGCAAGGTAGCGGTCAATGACACCTGCATTTGC 3694
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db GGCTCTCGGACCTTTACCTGGTCCAGGACGCGCGATGTCATTCCTGGTGGCGCGG 3754
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db GGTGATAGCAGGGGACGCTGCTGTCGCCCGGCCCATTTCTTCTACTTTGAAAGGCTCTCCG 3814
141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
3815 GGGGGTCCGCTGTGTGCCCCGGGGCAGCGGTGGGCATATTTAGGGCCGGGTGTGC 3874
161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
3875 ACCCGTGGAGTGGCTAAAGCGGTGGACTTTATCCCTGTGTGGAGAACCTAGAGACAACATG 3934
181 ArgSerProValPheSerAspAsnSerProAlaValProGlnSerTyrGlnVal 200
3935 AGTCCCCCGGTGTTCACGGATACTCTCTCCACAGTAGTGTGCCAGAGCTTCCAGGTG 3994
201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
3995 GCTCACTCCATGCTCCACAGCAGCAGCGGCAAAAGCACCAGGTCCCGGTGCATATGCA 4054
221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
4055 GCTCAGGGCTATAAGGTGTCTACTCAACCCCTCTGTGTGCTGCAACACTGGGCTTTGGT 4114
241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
4115 GCTTACATGTCGAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 4174
261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
4175 ACCACTGGCAGCCCCCATCAGTACTCCACTACGGCAAGTTCCTTGGCCGAGCGGGTGC 4234
281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
4235 TCGGGGGCGGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGGATGCAATCC 4294
301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
4295 ATCTTGGGCATCGGCACATGCTCTTGACCAAGCAGAGACTGCGGGGCGAGACTGGTGTG 4354
321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
4355 CTCGCCACCGCCACCCCTCCGGGCTCCGTCACGTGCCCCCATCCCAACATCAGGAGGTT 4414
341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
4415 GCTCTGTCCACCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTAATC 4474
361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
4475 AAGGGGGGAGACATCTCATCTTGTCTTCAAGAGAAAGTGCAGCAATCTCGCCGA 4534
381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
4535 AAGTGTGTCGCTTGGGCATCAATGCGGTGGCTACTACCGGGTCTTACGTCGTC 4594
401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
4595 ATCCCGACCGAGCGGATGTTGTCTGCTGGCAACCGATGCCCTCATGACCGGCTATACC 4654
421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
4655 GCGACTTCGATCGGTATAGCTGCATATCGTGTGTCCACGACAGCTGCATTTCAAC 4714
441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
4715 CTTGACCTTACCTTCCACATTGAGACAATCAGCTCCCCCGAGATGCTGTCTCCGCACT 4774
461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
4775 CAACGTTCGGGGGAGGAGTGGCAGGGGGAGCGCAGGATCTACAGATTTTGTGCACCGGG 4834
481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500
4835 GAGCGGCCCTCCGGCATGTTTCGCTCCGTCCTCTGTGAGTGTATGACGAGGCTGT 4894
501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
```

```
Db 4895 GCTTGGTATAGCTACACCCCGCAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4954
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGCTTCCTGGTGGCAGGACCACTTGAATTTGGGAGGGCGTCTTTACAGGCTC 5014
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5015 ACTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 5074
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 5075 CTGTAGCGTACCAAGCACCCTGTCCTAGGCGCTCAAGCCCTTCCCCCATCGTGGAC 5134
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCACCCCTCCATGGGCCACACCCCTGCTA 5194
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGCTGGCGCTGTTCAGATGAAATCACCCCTGACGACCCAGTCACCAATACATC 5254
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5255 ATGACATGCATGTCCGGCGACCTGGAGTGTCTACGAGCACCTGGGTGCTCGTTGGCGGC 5314
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTGCGCTGTTGGCGCGTATGTCTGTCACAGGCTCGTGGTCATAGTGGCAGG 5374
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTCTGCTGTCCGGAGCGGCAATCATACCTCACAGGGAAGTCTCTACCGAGAGTTC 5434
Qy 681 AspGluMetGluGluCys 686
Db 5435 GATGAGATGGAAGAGTGC 5452

RESULT 15
LOCUS AR166930 9379 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6284249.
ACCESSION AR166930
VERSION AR166930.1 GI:16243325
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9379)
AUTHORS Barban,V.
TITLE Fusion polypeptide having the C protein and E1 protein of hepatitis
C virus
JOURNAL Patent: US 6284249-A 1 04-SEP-2001;
FEATURES
source
Location/Qualifiers
1..9379
/mol_type="unassigned"
/organism="unknown"

ORIGIN
Alignment Scores:
Pred. No.: 1,71e-199 Length: 9379
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AR166930 (1-9379)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3395 CTGGCGCCCATCATCAGCGGTACGCCGACAGACAAAGGGGCGCTCTCTAGGGTGCATAATCACC 3454
```

```
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3455 AGCCTAACTGGCGGGACAAAACCAAGTGAGGGTGAGGTGCCAGATTGTGTCAACTGCT 3514
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3515 GCCCAAACTTCTCGGCAACGTCATCAATGGGGTGTCTGGACTGTCTACACGGGGCC 3574
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3575 GGAACGAGGACCATCGCGTCAACCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC 3634
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100
Db 3635 CAAGACCTTGTGGCTGGCGCTCCGCAAGTAGCCGCTCATTGACACCCCTGCACCTTGC 3694
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 3695 GGCTCTCGGACCTTTACCTGGTCACGAGGACCGCATGTCTATCCCGTCGCGCGCGG 3754
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3755 GTGATAGCAGGGGACGCTGTCTGCCCGCGGCCATTTCTACTTGAAGGCTCTCTCG 3814
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3815 GGGGCTCGCTGTGTGCCCGCGGGCACCGCTGGGCATATTTAGGGCGCGGTGTGC 3874
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3875 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGACAAACCATG 3934
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGluVal 200
Db 3935 AGTCCCCGGTGTTCACGATAACTCTCTCCACAGTAGTGGCCAGAGCTTCCAGGTG 3994
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 3995 GCTCACCTCCATGCTCCACAGCGAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA 4054
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4055 GCTCAGGGCTATAAGGTGTGTAGTACTCAACCCCTCTGTGTGTGCAACACCTGGGTGGT 4114
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4115 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACATTT 4174
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4175 ACCACTGGCAGCCCCATCATCGTACTCCACCTACGGCAAGTTCCTTGGCGAGCGGGTGC 4234
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGCGCTTATGACATAATAATTTGTACAGAGTGCACCTCCAGGATGCCACATCC 4294
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTTGGGCATCGGCACCTGTCTTGACCAAGCAGAGACTCGGGGGGCGAGACTGGTTGTG 4354
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4355 CTGGCACCGCCACCCCTCCGGGCTCGTCACTGTGCCCCCATCCCAACATCGAGGAGTT 4414
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4415 GCTCTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 4474
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4475 ARGGGGGGGAGACATCTCATCTTCTGTTCATTCAAAGAAAGAGTGCAGCAACTCGCGCA 4534
```


QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 4535 AAGCTGCTGGCATTTGGGCATCAATCGCGGCTACTACCGGGTCTTGACGTGCCGTC 4594
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 4595 ATCCCGACCGCGCGATGTTGTCGTGGGACCGGATGCCCTCATGACCGGCTATACC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 4655 GGGGACTTCGACTCGGTAGACTGCAATACGTGTGTCCACGACAGCTCGATTTCAGC 4714
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 4715 CTTGACCTTACCTTCACCATTTAGACAATCACGCTCCCCAGGATGCTCTCCCGCACT 4774
QY 461 GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 4775 CAACGTCCGGGAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCAGGG 4834
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 4835 GAGCGCCCTCCGGCATGTTCCGACTCGTCGTCCTCTGTGAGTGTATGACGAGGCTGT 4894
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 4895 GCTTGTGTATGAGTTCACGCCCGCGAGACTACAGTTAGGCTACGAGGTACATGAACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
DB 4955 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGGAGGGCGTCTTTACAGGCCTC 5014
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 5015 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
DB 5075 CTGGTAGCGTACCAAGCACCGTGTGCTAGGCTCAAGCCCTCCCCCATCGTGGGAC 5134
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 5135 CAGATGTGGAAGTGTGTTGATTGGCTCAAGCCCACTCCATGGGCCAACACCCCTGCTA 5194
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 5195 TACAGACTGGGCGCTGTTCAAGATGAAATCACCTGACGACCCAGCACCACCAATACATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
DB 5255 ATGACATGCATGTCGGCGGACCTGGAGGTGCTCAGAGCACCTGGGTGCTCGTTGGCGGC 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 5315 GTCTGTGGCTGTTTGGCGCGTATTGCTGTCAACAGGCTGCGTGGTATAGTGGGAGG 5374
QY 661 IleValLeuSerGlyLysProAlaIleLeuProAspArgGluValLeuTyrArgGluPhe 680
DB 5375 GTCGTCTTGTCCGGGAAGCCGCAATCATACCTGACAGGGAAGTCCCTCTACCGAGATTTC 5434
QY 681 AspGluMetGluGluCys 686
DB 5435 GATGAGATGGAGAGTGC 5452

XX (TRIP-) TRIPEP AB.
XX Sallberg M;
XX WPI; 2002-339446/37.
XX P-PSDB; AAE21837.
XX Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
XX absence of hepatitis C virus in a subject and for preparing a medicament
XX for treating hepatitis C virus infection.
XX Claim 1; Page 64-65; 90pp; English.
XX The present invention relates to novel hepatitis C virus (HCV) NS3/4A
XX proteins and their corresponding polynucleotides. NS3/4A sequences are
XX useful for identifying the presence or absence of HCV in a subject. They are
XX useful for preparing a medicament used for treating or preventing HCV
XX infection. Sequences of the invention are also used as vaccines. The
XX present sequence is a DNA encoding HCV NS3/4A protein
XX
XX Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 3,22e-245 Length: 2061
XX Score: 3618.00 Matches: 686
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-09-930-591-2 (1-686) x AAD34500 (1-2061)
XX
XX 1 MetAlaProIleThrAlaTrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
XX
XX 1 ATGGCGCTATACAGGCTATGCCAGACAGAGGGGCTTTGGGATCATATACACC 60
XX
XX 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
XX
XX 61 AGCTTGACCGCGCGGACAAACCAGGTGCGAGGGTGAGGTTCCAGATCGTGCACTGCT 120
XX
XX 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThrHisGlyAla 60
XX
XX 121 GCCCAGACTTCTTGGCAACCTGCATTAAACGGGGTGTGGTGGTCTGTACCATGGAGCC 180
XX
XX 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
XX
XX 181 GGACAAAGGACCATTTGGCTCACCTAAGGGTCCCTGTATCCAGATGTACCAATGTGGAC 240
XX
XX 81 GlnAspLeuValGlyTrpProAlaProGlnGlnGlyAlaArgSerLeuThrProCysThrCys 100
XX
XX 241 CNAAGCTCTGAGCTGGCCCGCTCCCAAGTGCCCGCTCATTAACACCATGCACTTGC 300
XX
XX 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
XX
XX 301 GGCTCTCTCGGACCTTTACCTGGTCACGAGGACGCGGATGTCATTCTGTGCGCGCACGG 360
XX
XX 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
XX
XX 361 GGTGATGGCAGGGGAGCGCTTTCGCCCGGCGCTATCTTACTTTGAAGGCTCTCTCG 420
XX
XX 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
XX
XX 421 GGAGGCGCTCTGTGTGCGCCCGCAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGC 480
XX
XX 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
XX
XX 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGGTAGAGACTTAGAGACAACCATG 540
XX
XX 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
XX
XX 541 AGGTCCCGGTGTCTCAGACAACTCTCTCCCAACAGAGTGCCCGCAGAGCTACCAAGTG 600

QY 201 AlaHisLeuHisAlaProThrThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 601 GCCACCTGTCATGCTCCACCGGCGGTAAAGACACCAAGGTCGCGCGCATACGCA 660
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 661 GCTCAGGGGTACAAAGGTGCTGCTCAACCCCTCGTGTGCTGCAACAATGGGCTTTGGT 720
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTCGAAGGCCCATGGATTGATCTTAACATCAGACTGGGTGAGACAAATT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
DB 781 ACTACTGGCAGCCGATCAGTATTCCACTACGGCAAGTTCCTGCCGACGCGGTGT 840
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 841 TCAGGGGGTGTCTATGACATAATAATTGTGACGAGTGCACCTCCACGGATGCAACATCC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 901 ATCTTGGGCAATTGGCACTGCTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACCTGTGCCCATCTTAACATCAGAGGAGTT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATCCCTTTGAAGCAATT 1080
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCACGAGCTCGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 1141 AMACTGGTCCGCTGGGCGTCAATGCGTGGCTTACTACCGCGCTTGATGTGCGTC 1200
QY 401 IleProThrSerGlyAspValValValAlaThrAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCCGACAGTGGTGACCTGTCGTGGCAACTGACGCCCTCATGACCGGCTTTACC 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 1261 GGGCACTTCGATTCGGTGATAGACTGCAACACGCTGTGTCCACGACAGCTCGACTTCAGC 1320
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCTTACCTTCCACTTGAGACAAATCAGCTTCCCGAGATGCTGTCTCCGCTACT 1380
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1381 CAACGTCGGGGTAGGACTGGCAGAGGAAGCCAGGCACTTACAGATTTGTGGCAGCGGG 1440
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrTrpAspAlaGlyCys 500
DB 1441 GAGCGTCTTCTGGCATGTTTGNACTCTGTGCTCTGTGAGTGTCTATGACGCGGGTGT 1500
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 1501 GCTTGTATGAGCTTACGCGCGCGGAGACCAAGATTAGGCTACGAGCATACATGAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
DB 1561 CCGGGACTTCCCGGTGTCAGACCACTCTGAATTTTGGAGGGGCTCTTTACGGGTCTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 1621 ACCCAGATAGACGCCCACTTCTCTATCCGAGACAAAGCAGAGTGGGGAACCTTCCCTAT 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580

Db 1691 CTGTAGCTACCAAGCACCGTGTGCGCTAGAGCTCAAGCCCTCCCTCCCTCGTGGAC 1740
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGGAAGTCTTGATCGGCTCAAGCCACCTCCATGGGCCAACCCTCTGCTA 1800
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TATAGACTGGGCGGTGTCCAGAAATGAAGTCAACCTGACGCCACCCAGTCAACAAATATATC 1860
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATGACATGATGTGCGGTGACCTGAGGTCTCAGAGTCTCAGAGTACCTGGTCTCGTTGGCGGC 1920
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTTCTGGCTGTTGGCCGCGTATGCTCTCAACAGCTCGGTGGTCTATAGTAGTAGG 1980
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 ATTGTCTGTCCGGAAGCGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTC 2040
Qy 681 AspGluMetGluGluCys 686
Db 2041 GATGAATGGAAGAGTGC 2058

RESULT 2

AAD31767
ID AAD31767 standard; DNA; 2061 BP.
XX
AC AAD31767;
XX
DT 18-JUN-2002 (first entry)
XX
DE Hepatitis C virus (HCV) NS3/4A DNA coding region.
XX
DE Hepatitis C virus; HCV infection; virucide; fungicide; antibiotic;
KW cytosolic; immunostimulant; vaccine; ribavirin; immune response; cancer;
KW ds.
XX
OS Hepatitis C virus.
XX
FH Location/Qualifiers
FT 1. .2061
FT /*tag= a
FT /product= "HCV NS3/4A protein"
XX
PN WO200213855-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001808.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX
PA (TRIP-) TRIPEP AB.
XX
PI Sallberg M, Hultgren C;
XX
PI WPI; 2002-241837/29.
XX
PI P-PSDB; AAE19900.
XX
PT Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
PS Claim 1; Page 94-95; 120pp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of

CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is HCV NS3/4A DNA coding region
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
Alignment Scores:
Pred No.: 3.22e-245 Length: 2061
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-930-591-2 (1-686) x AAD31767 (1-2061)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1 ATGGCGCTATCACGGCTATGCCAGCAGACAAGGGGCTTTTGGGATGCATAATCACC 60
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGluValGlnIleValSerThrAla 40
Db 61 AGCTTGACCGCGCGGACAAAACCCAGGTGGAGGTGAGGTTCAGATCGTCAACTGCT 120
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 121 GCCAGACTTCTTGGCAACTGCTTAACGGGGTGTGGACTGTCTACATGGAGGCC 180
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 181 GGAACAAGGACCATTTGCTCACCTAAGGTCCTGTATCCAGATGTACCAATGTGGAC 240
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 241 CAAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCGCTCATTAACACCATGACATTGC 300
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 301 GGCTCTCGGACCTTTTACCTGGTCACGAGGCACGCCGATGTCATCTCTGCGCGGACGG 360
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 361 GGTGATGGCAGGGGCGAGCTGCTTTTCCGCCCGGCTATCTCTTACTTGAAGGCTCTCG 420
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 421 GGAGGCCCTCTGCTGTGCCCGCGCAGGACATGCCGTAGGCATATTTCAGAGCGCGGTATGC 480
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 481 ACCGTGGAGTGGCTAAGGCGGTGGACTTATCCCTCCGTAGAGAGCTTAGAGACACCAATG 540
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 541 AGTCCCGGTGTCTCAGACAACCTCTCCCCACGACGAGTCCCGCAGAGTACCAAGTG 600
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 601 GCCCACCTGCATGCTCCCGCGGCGGTAAAGACCAAGGTCCCGCGCGCATACGCA 660
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 661 GCTCAGGGCTACAGGTGCTGGTGTCAACCCCTCCGTGTCTGCAACAATGGGCTTTGGT 720
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 721 GCTTACATGTCCAAGGCCCATGGATTGATCTTAACATCAGGACTGGGGTGAGGCAATT 780
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280

Db 781 ACTACTGGCAGCCGATCAGCTATTCACCTACGGCAAGTTCTTGGCGAGCGGGTGT 840
Qy 281 SerGlyAlaValAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 841 TCAGGGGGTCTTATGACATATATTTTGTGACAGTGCACCTCCACGGATGCAACATCC 900
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 901 ATCTTGGCATTTGGCATCTCTTACCAAGCAGAGACCGGGGGGAGACTGACTGTG 960
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 961 CTGCCACCGCTACCCCTCGGGCTCGCTCAGTGTGCCCTTACATCAGAGAGTT 1020
Qy 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeuGluAlaIle 360
Db 1021 GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATTTCCCTTGAAGCAAT 1080
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 1081 AAGGGGGGAGACATCTCTCTCTCCACTCAAAGAAAGTGGCAGAGCTCGCGCA 1140
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaValAlaValAlaValAlaVal 400
Db 1141 AAATCTGTCGGTGGGGCTCAATGCGCGTCTTACTACCGCGGCTTGTATGTGTCGTC 1200
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 1201 ATCCGACAGTGTGTGAGTGTGCTGTGGCACTGACGCCCTCATGACGGCTTTACC 1260
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 1261 GCGACTTCGATTCGGGTAGACTGCAACAGTGTGTCAACCAGACAGTCTGACTCAGC 1320
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTTGACCCCTACCTTCACTATGACAAATCAGCTTCCCGAGATGCTGTCTCCGCTACT 1380
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleThrArgPheValAlaProGly 480
Db 1381 CAAGTGGGGTAGGACTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Qy 481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysThrAspAlaGlyCys 500
Db 1441 GAGCGTCTCTGCGATGTTGACTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Qy 501 AlaTrpThrGluLeuThrProAlaGluThrThrValArgLeuArgAlaThrMetAsnThr 520
Db 1501 GCTTGGTATGAGCTTACGCGCGCGGAGACACACAGTTAGGCTACGAGCATACATGAAC 1560
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 1561 CCGGGACTTCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGTCTTTACGGTCTC 1620
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProThr 560
Db 1621 ACCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
Qy 561 LeuValAlaThrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580
Db 1681 CTGGTAGCGTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGAC 1740
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGGAGTGTGTGATCCGTCTCAAGCCACCTCTCATGGGCAACACCTCTGCTA 1800
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysThrIle 620
Db 1801 TATAGACTGGGGCTGTCTCAGATGAATGAATCAGCTCAGCCAGCCAGTCAACAGTATATC 1860
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGly 640
Db 1861 ATGACATGTATGTGGCTGACCTGGAGGTGCTCAGCAGTACCTGGGTGCTCGTGGCGC 1920

Qy 641 ValLeuAlaLeuAlaAlaValCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTTCTGGCTGCTTGGCCGGTATTGCCCTATCCACAGGCTGCGTGCATAGTAGTAGG 1980
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuThrArgGluPhe 680
Db 1981 ATTGTCTTGTCCGAAAGCCGGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTC 2040
Qy 681 AspGluMetGluGluCys 686
Db 2041 GATGAATGGAAGAGTGC 2058
RESULT 3
AAD60868
ID AAD60868 standard; DNA; 2061 BP.
XX
AC AAD60868;
XX
DT 15-JAN-2004 (first entry)
XX
DE Hepatitis C virus NS3/4A DNA.
XX
KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 1..2061
FT /tag= a
FT /product= "Hepatitis C virus protein"
XX
PN US2002136740-A1.
XX
PD 26-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00929955.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
XX
PA (SALL/) SALLBERG M.
PA (HULT/) HULTGREN C.
XX
PI Sallberg M, Hultgren C;
XX
DR WPI; 2003-764978/72.
DR P-PSDB; ABW00351.
XX
PT Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
PS Claim 1; Page 60-61; Opp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS3/4A DNA
XX
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,22e-245 Length: 2061
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-930-591-2 (1-686) x AAD60868 (1-2061)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1 ATGGGGCTATCATCGGCTATGCCAGACAGACAGAGGGGCTTTGGGATGCATATCACC 60
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Db 61 AGCTTGACCGCGCGGACAAACACAGGTGGAGGGTGAGGTTTCAGATCGTCAACTGCT 120
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 121 GCCCAGACTTTCTTGGCAACCTGCATTAAACGGGGTGTGTGAGATGCTACATGGAGCC 180
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 181 GGRACAGGACCATTCGCTACCTTAAGGGTCTCTGTATCCAGATGTACCAATGTGGAC 240
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 241 CAAGACCTCGTAGCTGGCGCGCTCCCAAGGTGCGCGCTCATTAACACCATGCACTGC 300
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 301 GGCTCTCGGACCTTACCTGGTCAAGGACACCGCGATGTCATTCTGTGCGCGGACGG 360
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer 140
Db 361 GGTGATGGCAGGGAGCGCTCTTTCGCCCGGCTATCTCTTACTTGAAGGCTCTCTCG 420
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 421 GGAGGCTCTGCTGTGCGCGCAGGACATGCCGTAGGCATATTACAGAGCGCGGTATGC 480
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuLeuThrThrMet 180
Db 481 ACCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGGTAGAGCTTAGAGACAACCATG 540
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 541 AGTCCCCGGGTGTTCTCAGACACTCTCTCCACACAGCAGTGCCTCCAGAGTCAAGTG 600
Qy 201 AlaHisLeuHisAlaProThrGlySetGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 601 GCCCACCTGCTGCTCCACCGCAGCGGTAAAGACACCAAGGTCCCGCGCATACGCA 660
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 661 GCTCAGGGCTACAAAGGTGCTGGTGTCAACCCCTCCGTGTGTCGAACAATGGGCTTTGGT 720
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 721 GCTTACATGTCACAGGCCCATGGGATGATCCTAACATCAGAGTGGGGTGAGACAATT 780
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 781 ACTACTGGCAGCCGATCAGTATTCCACCTACCGCAAGTTCCTTGGCGAGCGGGGTGT 840
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 841 TCAGGGGGTGCTTATGACATAATAATTGTGACGAGTGCCTCCAGGATGCAACATCC 900
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 901 ATCTGGGCATTTGGCACTGCTCTTGACCAAGCAGAGACCGCGGGGGGAGACTCACTGTG 960
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 961 CTGCGCACCGTACCCCTCCGGGCTCCGTCTACTGTGCCCCATCTTAACTCCAGAGAGTT 1020
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
|||||

RESULT 4
ADG47658
ID ADG47658 standard; DNA; 2061 BP.

Db 1021 GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTGAAGCAATT 1080
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 1081 AAGGGGGGAGACATCTCATCTCTCTCCACTCAAAAGAAAGTGCAGACGAGCTCGCGCA 1140
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 1141 AAATGTGTCGCTTGGCGGTCAATGCGGTGCTTACTACCGCGCTTGTGTCTCGCTC 1200
Qy 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 1201 ATCCGACCAAGTGTGTGAGCTTGTCTGTGCAACTGACGCCCTCATGACCGGCTTACC 1260
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 1261 GCGGACTTCGATTCGGGTGATAGACTGCAACACGCTGTGTCAACCCAGACAGCTTCAGC 1320
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTTGACCTTACCTTCAACATTGACAAATCAGCTTCCCGAGGATGCTCTCCCGTACT 1380
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 1381 CAACGTGGGGTAGGACTGGCAGAGGAGAGCCAGGCATCTACAGATTTGTGGCACCGGG 1440
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 1441 GAGGCTCTTCTGGCATGTTTGTACTCTCTCTCTCTCGAGTGTATGACGCGGGTGT 1500
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 1501 GCTTGGTATGAGCTTACGCCCGCGAGACACACAGTACAGGTACGAGCATACATGAACACC 1560
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 1561 CCGGGACTTCCCGGTGCCAAGACCATCTTGAATTTTGGGAGGGCGCTTTTACGGGCTC 1620
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 1621 ACCACATAGACGCCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerThrAsp 580
Db 1681 CTGCTAGCTACCAAGCACCGTGTGCGGTAGAGCTCAAGCCCTCCCGCTCGTGGGAC 1740
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGGGAAGTCTTGTATCCCGTCTCAAGCCCACTTCCATGGGCCCAACACCTCTGCTA 1800
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TATAGACTGGGCGCTGTCCAGATGAAGTCACTCCCTGACGACCCAGCTCACCAGATATATC 1860
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrThrValLeuValGlyGly 640
Db 1861 ATGACATGTATGTGCGGTGACCTGGAGTGTGTCACGAGTACCTGGGTGCTCGTGGCGGC 1920
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTTCTGCTCTTGGCGCGGTATTGCTTATCCACAGGCTGCGGTGCTATAGTAGGTAGG 1980
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 ATTGTCTTGTCCGGAAGCCGGCATATATCCCGACAGGAGAGTCTTCTACCGGGAGTTC 2040
Qy 681 AspGluMetGluGluCys 686
|||||
Db 2041 GATGAATGGAAGAGTGC 2058
|||||

XX AC ADG47658;
 XX DT 11-MAR-2004 (first entry)
 XX DE HCV NS3/4A domain DNA.
 XX KW ds; gene; immunogen; hepatitis C virus; HCV infection; vaccine.
 XX OS Hepatitis C virus.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..2061
 XX FT /*tag= a
 XX FT /product= "NS3/4A domain"
 XX XX US2003206919-A1.
 XX XX 06-NOV-2003.
 XX XX 26-NOV-2002; 2002US-00307047.
 XX XX 17-AUG-2000; 2000US-0225767P.
 XX XX 29-AUG-2000; 2000US-0229175P.
 XX XX 15-AUG-2001; 2001US-00929955.
 XX XX 15-AUG-2001; 2001US-00930591.
 XX PA (SALL/) SALLBERG M.
 XX XX Sallberg M;
 XX PI WPI; 2004-051480/05.
 XX DR P-PSDB; ADG47659.
 XX DR
 XX PT New purified or isolated nucleic acid useful for enhancing an immune
 XX PT response to a hepatitis C antigen comprises specific nucleotide sequences
 XX PT and the amino acid sequences.
 XX XX
 XX PS Example 1; SEQ ID NO 1; 83pp; English.
 XX CC The invention relates to a purified or isolated nucleic acid. The
 XX CC peptides are useful as immunogens for the treatment and prevention of
 XX CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
 XX CC The nucleic acid and the peptide enhance an immune response to a
 XX CC hepatitis C antigen and are potent immunogens. The present sequence is
 XX CC used in the exemplification of the invention.
 XX XX
 XX SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3 22e-245 Length: 2061
 Score: 3618.00 Matches: 686
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0
 US-09-930-591-2 (1-686) x ADG47658 (1-2061)
 QY 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1 ATGGCGCTATCATCGGCTATGCCAGACAGAGGGGCTTTGGGATGCATAATCACC 60
 QY 21 SerIleuThrGlyArgAspLysAenGlnValGluGlyValGlnIleValSerThrAla 40
 DB 61 AGCTTGACCGCGCGGACAAAACCAAGGTGGAGGGTGAGGTTTCAGATCGTCAACTGCT 120
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAenGlyValCysTrpThrValTyrHisGlyAla 60
 DB 121 GCCCAGACTTCTTGGCACTTCATTAACGGGGTGTGTGGACTGTCTACCATGAGGC 180
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

DB 181 GGAACAAGGACCATTGCGTCACTAAGGGTCTCTTTATCCAGATGTACACCAATGTGGAC 240
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 241 CAAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACACATGCACTTGC 300
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 301 GGCTCTCTCGACCTTTACCTGTGTACAGAGCAGCGCGATGTATTCTCTGTGGCCGACGG 360
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 DB 361 GGTGATGGCAGGGCAGCCTGCTTTCCGCCCGCGCTATCTTACTTTAAAGGCTCCTCG 420
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 421 GGAGGCCCTCTGCTGTGCCCGCAGACATGCCGTAGGCATATTTCAGACCGCGGTATGC 480
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 DB 481 ACCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGTAGAGAGCTTAGACACACCATG 540
 QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
 DB 541 AGGTCCCGGTGTCTCTCAGACAACTCTCCCCACAGCAGTGTCCCGCAGAGCTACCAAGTG 600
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 601 GCCCACCCTGCATGCTCCACCGCAGCGGTAAAGACCAACAGGTCCCGCCGCATACGCA 660
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAnProSerValAlaAlaThrMetGlyPheGly 240
 DB 661 GCTCAGGCTACAGGTGTGTGTCTCAACCCCTCGTGTCTGCAACATGGGCTTGTGT 720
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAenIleArgThrGlyValArgThrIle 260
 DB 721 GCTTACATGTCCAAGGCCCATGGATTGATCTCTAAACATCAGGACTGGGTGAGGACAT 780
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 781 ACTACTGGCAGCGCCCATCATCACGATTCCACTACGGCAAGTTCCTTCCCGCAGCGGGTGT 840
 QY 281 SerGlyGlyValTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 841 TCAGGGGGTGTCTATGACATATATTTGTGACGAGTGCCTCCACGGATGCAACATCC 900
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgIleThrVal 320
 DB 901 ATCTTGGCATTTGGCACTGTCTTGACCAAGCAGAGACCGCGGGGCGAGACTGACTGTG 960
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCATCTGTGCCCATCTTAACATCAGGAGGT 1020
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 DB 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGCCAAGGCTATTCCTTGAAGCAAT 1080
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 DB 1081 AAGGGGGGAGAGACATCTCATCTCTCTGCCACTCAAAAGAGAGTGCAGAGCTCGCGCA 1140
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 1141 AAACGTGTCGCTGTGGCGCTCAATCGGTGGCTTACTACCGGGCTTGTATGTGTCGGTC 1200
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 1201 ATCCCGACCGAGTGTGCGTCTGTGTGGCACTGACGCCCTCATGACCGCTTTACC 1260
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 1261 GCGGACTTTCGATTTCGTGTAGACTGCAACACGTGTGTCCACGACAGTCAGCTTCAGC 1320

Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTTGACCCCTAGCTTACCATTGAGACAATCAGCTTCCAGGATGCTGTCTCCGCTACT 1380
Qy 461 GlnArgArgGlyArgThrGlyArgGlyPspRodGlyIleTyrArgPheValAlaProGly 480
Db 1381 CAACGTCGGGGTAGGACTGGCAGAGGGAAGCAGGCATCTACAGATTTGTGGCACCAGGG 1440
Qy 481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 1441 GAGCGTCTCTTGGCATGTTGACTCGTCTGTCTCGGAGTGTATGACGCGGGTGT 1500
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 1501 GCTTGGTATGACTTACGCCCGCCGAGACACACAGTTAGGCTACGAGCATACATGAACACC 1560
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 1561 CCGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACGGGTCTC 1620
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 1621 ACCACATAGAGCCCATCTCTATCCAGACAAGACAGAGTGGGAAACCTTCCCTAT 1680
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 1681 CTGGTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCTCCCGTGTGGGAC 1740
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGGAAGTGTGTCGTCGTCCTCAAGCCCAACCTTCCATGGGCCCAACCTCTGCTA 1800
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TATGACTGGCGCGTGTCCAGAAAGTACCTGACCGACCCAGTCCACCAAGTATATC 1860
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATGACATGATGTGCGCTGACCTGGAGGTGTCACAGTACCTGGGTGCTGTTGGCGGC 1920
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTTCTGGCTGCTTTGGCGCGCTATTGCTATCCACAGGCTGCGTGGTCAATAGTAGTAGG 1980
Qy 661 IleValLeuSerGlyLysProAlaTyrIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 ATTGTCTTGTCCGGAAGCCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTC 2040
Qy 681 AspGluMetGluGluCys 686
Db 2041 GATGAATGGAGAGTGC 2058

RESULT 5

ADG47692

ID ADG47692 standard; DNA; 2078 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

PD 06-NOV-2003.

XX 26-NOV-2002; 2002US-00307047.

XX 17-AUG-2000; 2000US-0225767P.

XX 29-AUG-2000; 2000US-0229175P.

XX 15-AUG-2001; 2001US-00929955.

XX 15-AUG-2001; 2001US-00930591.

XX (SALL/) SALLBERG M.

XX Sallberg M;

XX WPI; 2004-051480/05.

XX P-PSDB; ADG47693.

XX New purified or isolated nucleic acid useful for enhancing an immune

XX response to a hepatitis C antigen comprises specific nucleotide sequences

XX and the amino acid sequences.

XX Claim 2; SEQ ID NO 35; 83pp; English.

XX The invention relates to a purified or isolated nucleic acid. The

XX peptides are useful as immunogens for the treatment and prevention of

XX hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.

XX The nucleic acid and the peptide enhance an immune response to a

XX hepatitis C antigen and are potent immunogens. The present sequence is

XX used in the exemplification of the invention.

XX SQ Sequence 2078 BP; 351 A; 834 C; 630 G; 263 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,26e-245 Length: 2078

Score: 3618.00 Matches: 686

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

US-09-930-591-2 (1-686) x ADG47692 (1-2078)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleThr 20

Db 12 ATGGCCCCATCAGCGCTACGCCAGACAGCCGCGCTGTGGCTGCATCATCACC 71

Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluValGlnIleValSerThrAla 40

Db 72 AGCTGTACCGCGCGCGCAAGAACCCAGGTGGAGGCGAGGTGCAGATCGTGAGCACCGCC 131

Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60

Db 132 GCCAGACCTTCTGGCCACCTGCATCAACGGGTGTGTGGACCGTGTACACCGCGCC 191

Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

Db 192 GGCACCGCGCACCATCGCCAGCGCCCAAGGGCCCCGTGATCCAGATGTACCAACGTGGAC 251

Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100

Db 252 CAGGACCTGTGGGCTGGCGCGCCCGCCAGGGCGCGCGAGCTGACCCCTGTACACCTGC 311

Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120

Db 312 GGCAGACGCGACTGTACCTGTGTGACCCCGCCAGCGCGCGAGTGTATCCCGCGCGCCGC 371

Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140

Db 372 GGCAGCGCGCGCGAGCTGTGTGAGCCCCCGCCCATCATCAGTACCTGAGAGGCGAGCAGC 431

Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160

Db 432 GGCAGCGCGCGCGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 491

```
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 492 ACCCGCGCGTGGCCAGGCGGTGACTTCATCCCGTGGAGAGCTGGAGACCACTG 551
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 552 CGCAGCCCGCGTGTTCAGCGCAACAGCAGCGCCCGCCCGCGTGGCCAGAGCTACCAAGTG 611
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaThrAla 220
DB 612 GCCCACTGCGCGCCCGCCCGCGGCGGCAAGAGCACCAGGTGGCCCGCGCTACGCC 671
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 672 GCCCAGGGCTACAAAGTGTGTGCTGAACCCCGAGCGTGGCCGCCACCATGGGCTTCGGC 731
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 732 GCCTACATGAGCAAGGCCCGCCCGCATCGACCCCAACATCCGCGCCCGCGTGGCGCACATC 791
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 792 ACACCGCGAGCCCATCACTACAGCACTACGGCAAGTTCCTGGCCGACGGCGGTGC 851
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 852 AGCGCGCGCGCTACGACATCATCTCGCAGAGTGGCCACAGCAGCGCCACCGC 911
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 912 ATCCTGGGCATCGGCACCGTGTGGACAGCGCGCAGACCGCGCGCGCTGACCGTG 971
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 972 CTGGCCACCGCCACCCCGCGGAGCGTACCGTGGCCCGCCACCCCAACATCGAGGAGGTG 1031
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 1032 GCCTGAGCACCACCGCGGAGATCCCTTCTACGGCAAGGCCATCCCTCGGAGGCCATC 1091
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 1092 AAGCGCGCGCGCCACTGTATCTTCGCCACAGCAAGAAGTGGCAGCAGCTGGCGCC 1151
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 1152 AAGCTGTGTCCCTGGCGGTGAACCGCGTGGCTACTACCGCGCGCTGGACGTGACGGTG 1211
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 1212 ATCCCAACAGCGCGAGCGTGTGTGTGTGGCCACCGAGCGCTGTATGACCGGCTTCACC 1271
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 1272 GCGGACTTCGACAGCGGTGATCGATCGCAACACCTGCGTGACCCAGACCGCTGGACTTCAGC 1331
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1332 CTGGACCCCACTTCACATCGAGACCATCACTCCCGCCAGGACCGCGTGAGCCGACCC 1391
QY 461 GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1392 CAGCGCGCGCGCGCAGCGCGCGCGCAAGCCCGCATCTACCGCTTCGTGGCGCCCGCGC 1451
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 1452 GAGCGCGCCCGCGCGCATGTTCGACAGCAGCGGTGCTGTGGAGTGTACGACCGCGCGTGC 1511
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 1512 GCCTGTGTACAGTGTACCCCGCGGAGACCATCCGTCGCGCTGCGCGCTACATGAACACC 1571
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
```

```
DB 1572 CCGGCGCTGCGCGTGTGGCCAGGACCACTGGAGTCTTGGAGGGCGTGTTCACCGGCGCTG 1631
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 1632 ACCCATCGAGCGCCACTTCTCTGAGCCAGACCAAGCAGAGCGCGGAGAACCTGCCCTAC 1691
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
DB 1692 CTGTGGCGCTTACAGGCGCACCGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAC 1751
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 1752 CAGATGTGGAAGTGCCTGATCCGCTGGAAGCCACCCCTGCACGGCGCCACCGCGCTGCTG 1811
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 1812 TACCGCGCTGGCGCGCGTGCAGAACGAGGTGACCTGACCCCGCGGTGACCAAGTACATC 1871
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGly 640
DB 1872 ATGACCTGATGAGCGCGCGCTGAGGTGTGTGACAGACCTGGGTGCTGGTGGCGCGC 1931
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 1932 GTGCTGGCGCGCTGGCGCGCTACTGCTGAGCAGCGCGCTGCTGTGTGATCGTGGCGCGC 1991
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 1992 ATCGTGTCTGAGCGGCAAGCCCGCATCATCCCGCGAGGTGCTGTACCGCGAGTTTC 2051
QY 681 AspGluMetGluGluCys 686
DB 2052 GACGAGATGAGGAGTGC 2069
RESULT 6
AAN92106
ID AAN92106 standard; DNA; 7310 BP.
XX AC AAN92106;
XX 09-SEP-2004 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX DE Combined ORFs of the HCV cDNAs from clones K9-1 through 15e.
XX KW Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH; ss.
XX OS Hepatitis C virus.
XX Unidentified.
FH Key Location/Qualifiers
FT CDS 3..7310
FT /*tag= a
XX EP318216-A.
XX PD 31-MAY-1989.
XX PF 18-NOV-1988; 88EP-00310922.
XX PR 18-NOV-1987; 87US-00122714.
XX PR 30-DEC-1987; 87US-00139886.
XX PR 26-FEB-1988; 88US-00161072.
XX PR 06-MAY-1988; 88US-00191263.
XX PR 26-OCT-1988; 88US-00263584.
XX PR 14-NOV-1988; 88US-00271450.
XX PA (CHIR ) CHIRON CORP.
XX PA (CHIR ) CHIRON CORP.
XX PI Houghton M, Choo QL, Kuo G;
```


Db 3408 CTGTAGCGTACCAGCCACCGTGTGGCTAGGCTCAAGCCCTCCCTCCCATCGTGGGAC 3467
QY 581 GlnMetTrpLysCysLeuIleAsgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGGAGAGTTTGTATGGCTCAAGCCACCTCCATGGGCCAACCCCTGCTA 3527
QY 601 TtArgLeuGlyAlaValGlnAsgGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3528 TACAGACTGGCGCTGTTCAAGATGAATCACCCTGACGACCCAGTCACCAATACATC 3587
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3588 ATGACATGCATGTGGCGGACCTGGAGTGTCTACGAGCACCTGGGTGCTCTGGCGGC 3647
QY 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTGTGCTTTGGCGCGATTTGCTGTCAACAGCTGCGTGGTCAATAGTGGGAGG 3707
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTGCTCTTGTCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGATTTC 3767
QY 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785

RESULT 7

AAAN90336
ID AAAN90336 standard; DNA; 7310 BP.

AC AAAN90336;

XX 25-MAR-2003 (revised)

DT 19-JUL-2001 (revised)

DT 01-NOV-1989 (first entry)

XX Composite hepatitis C virus (HCV) cDNA.

XX Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.

XX Pan troglodytes.

XX GB2212511-A.

XX 26-JUL-1989.

PF 18-NOV-1988; 88GB-00027024.

XX 18-NOV-1987; 87US-00122714.

PR 30-DEC-1987; 87US-00139886.

PR 26-FEB-1988; 88US-00161072.

PR 26-OCT-1988; 88US-00263584.

XX (CHIR) CHIRON CORP.

PA Houghton M, Choo QL, Kuo G;

XX WPI; 1989-215054/30.

DR P-PSDB; AAP90288.

XX Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.

XX Disclosure; Fig 47; 30pp; English.

XX The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by
CC aligning clones k9-1 through 15e in 5'-3' direction. The cDNA encodes
CC antigens which react with antibodies in patients with non-A non-B
CC hepatitis (NANBH). The cDNA can be used to design probes, or to
CC synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to
CC raise antibodies for immunoassay or treatment, or to produce vaccines.
CC See also AAP90288, and AAAN90303-35. (N.B. This record was resubmitted to

CC correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR
CC field.)
XX Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 U; 0 Other;
SQ Alignment Scores:
Pred. No.: 1-88e-241 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0
US-09-930-591-2 (1-686) x AAAN90336 (1-7310)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 CTGGCCCATCATCGGCGTACGCCAGCAGACAAAGGGGCTCTTAGGGTGCAATACCC 1787
QY 21 SerLeuThrGlyArgAspLysAsgGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 1788 AGCCTAACTGGCGGGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCACTGCT 1847
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsgGlyValCysTrpThrValTyrHisGlyAla 60
Db 1848 GCCAAACCTTCTGGCAACGTGCATCAATGGGTGTGTGGACTGTCTACCAAGGGGCC 1907
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1908 GGAACGAGACCATCGCGTCACCAAGGCTCTGTCTATCCAGATGTATACCAATGTAGAC 1967
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1968 CAAAGACCTTGTGGGTGGCGCGCTCGCAAGTAGTACCGCTCATTTGACACCTTGCATTGC 2027
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 2028 GGTCTCTCGGACCTTTACCTGTGTACGAGGACGCGGATGTCTATCCCTGCGCGCGCG 2087
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 2088 GGTGTATAGCAGGGGCGAGCTGTGTGCGCCCGCGCCATTTCTTACTTTGAAAGCTCTCG 2147
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 2148 GGGGTCTCGCTGTGTGCGCGCGGGGACGCGGTGGGATATTTAGGCGCGGGTGTGC 2207
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
Db 2208 ACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGACCAACCATG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGGTCCCGGTGTTCACGGATAACTCTCTCACAGTAGTGTCCCGAGAGTTCACAGGTG 2327
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 2328 GTTCACTTCATGCTCCACAGCGGCGGCAAAAGCACCAGGTCCCGCTGCATATGCA 2387
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTGTGTGTCGAACACTGGGCTTGT 2447
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCGAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 2507
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCATGGGAGCGCCCATCAGTACTCCACCTACGCGCAAGTTCCTTCCGCGCGGGTGC 2567
QY 281 SerGlyGlyAlaTyrAspIleIleLeuCysAspGluCysHisSerThrAspAlaThrSer 300

```
Db 2568 TCGGGGGCGCTTATGACATATAATTGTGACGAGTCCACTCCACGATGCCACATCC 2627
Qy 301 lIeLeuGlylIeGlyThrValLeuAspGlnAlaGluThrAlaGlyValAArgLeuThrVal 320
Db 2628 ATCTTGGGCATCGGCACCTGCTTGTGACCAAGCAGAGACTGCGGGGCGAGACTGGTTGTG 2687
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTCGCCACCGCACCCCTCCGGGCTCCGCTACTGTGCCCCATCCCAACATCGAGAGGTT 2747
Qy 341 AlaleuSerThrThrGlyGluLeuProPheThrGlyLysAlaIleProLeuGluAlaIle 360
Db 2748 GCTCTGTCCACACCGAGAGATCCCTTTTACGCAAGGCTATCCCTCGAGTAATC 2807
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerTysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAGAGAAAGTGCAGCAATCGCGCA 2867
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaThrValArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGTGTGCGATTTGGGATCAATGCCGTGGCTACTACCGCGGTCTTTGACGTGTCCGTC 2927
Qy 401 lIeProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCCGACCGCGGGGATGTTGTCTGTGGCAACCGATGCCCTCATGCCGGCTATACC 2987
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTGCATTTACG 3047
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCCCTACCTTCAACATTGAGACAATACGCTCCCGCCAGGATGCTGTCTCCGCACT 3107
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleThrArgPheValAlaProGly 480
Db 3108 CAACGTCCGGGCGAGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGG 3167
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCGGCATGTTGCACTCGTCGCTCTGTGAGTGTATGACGCGAGCTGT 3227
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 3228 GCTTGGTATGACTCAGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 3287
Qy 521 ProGlyLeuProValCysGlnAspHisIleGluPheThrGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCGCTGTCAGGACCATCTTGAATTTTGGGAGGGCGCTTTTACAGGCTC 3347
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCACTTCTATCCAGACAAGCAGAGTGGGGAGAACCTTCTTAC 3407
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 3408 CTGTGTAGCGTACCAAGCCACCGTGTGCTAGGCTCAAGGCCCTCCCCCATCGTGGAC 3467
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGGAAGTGTGTGATTGCTCCCTCAAGCCACCTCCATGGGCCAACCCCTGTGA 3527
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3528 TACAGACTGGGGCGCTTTCAGATGAAATCACCTTGACGCCACCCAGTCCCAATATCATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3588 ATGACATGATGTCGCGCCGACCTGGAGGTCTCGTACAGACACCTGGGTGCTCGTTGGCGGC 3647
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTCTGGCTGCTTTGGCGCGTATTGCTGTCAACAGGCTGCTGCTCATAGTGGGCAGG 3707
```

```
Qy 661 lIeValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTCTGTCTGTTCGGGAAGCCGCAATCATATCCTACAGGGAAGTCTCTTACCGAGGTTTC 3767
Qy 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785
RESULT 8
ID AAQ98221 standard; cDNA to mRNA; 7310 BP.
XX AAQ98221;
AC AAQ98221;
DT 25-MAR-2003 (revised)
DT 15-AUG-1996 (first entry)
XX Hepatitis C virus clone genome.
DE Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;
KW antibodies; immunoprophylaxis; sera; serum; ds.
XX Hepatitis C virus.
OS US5443965-A.
PN 22-AUG-1995.
XX 05-APR-1991; 91US-00681703.
XX 06-APR-1990; 90US-00505611.
PR 09-OCT-1990; 90US-00594854.
XX (GENE-) GENELABS INC.
XX Mosekli R, Reyes GR, Kim JP;
XX WPI; 1995-302120/39.
DR New nucleic acids encoding hepatitis C virus antigens - used to develop
PT prods. for detection of HCV-infected sera and prodn. of vaccines and anti
PT -HCV antibodies.
PS Example 4; Fig 11; 71pp; English.
XX Hepatitis C virus (HCV) antigens can be used for detecting HCV infected
CC sera and individuals infected with HCV. They can also be used in an anti-
CC HCV vaccine or for the production of anti-HCV antibodies which can be
CC used for passive immunoprophylaxis. The antigens consistently identify
CC more HCV positive serum samples with a high degree of specificity. See
CC AAQ98202-14 and AAR81939-51. (Updated on 25-MAR-2003 to correct PF
CC field.) (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 7310 BP; 1494 A; 2217 C; 2060 G; 1539 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,88e-241 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 2 Gaps: 0
US-09-930-591-2 (1-686) x AAQ98221 (1-7310)
Qy 11 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 CTGGCGCCCATCCGCGGTACGCCAGCAGCAAGGGCCCTCTTAGGTGCATATCACC 1787
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 1788 AGCCTTAAGTCCCGGGCAAAAACCAAGTGGAGGGTGGAGTCCAGATTGTGTCAACTGCT 1847
```

```
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyVala 60
Db 1848 GCCCAAACTTCCTGGCAACGTCATCAATGGGGTGTCTGGACTGTCTACACAGGGGCC 1907

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1908 GGAACAGAGACCATCGCGTCAACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 1967

QY 81 GluAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1968 CAGACCTTGTGGGCTGGCCGCTCCGCAAGGTAGCGCTCATTCAGACCCCTGCATTCG 2027

QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 2028 GGCTCCTCGGACCTTACCTGGTACGAGGACGCCGATGTCAATCCCGTGGCGGGCGG 2087

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleuLysGlySerSer 140
Db 2088 GGTGATAGCAGGGGACGCTGTCTGCGCCCGGCCCATTTCTTACTTTGAAAGGCTCCTCG 2147

QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaIleValCys 160
Db 2148 GGGGGTCCGCTGTGTGCCCGGGGGGACGCGGTGGGCATATTAGGGCCGGGTGTGC 2207

QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 2208 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 2267

QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGGTCCCGGTGTTCACGGATAACTCTCTCCACCATAGTAGTGGCCACAGAGCTTCCAGGTG 2327

QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 2328 GCTCACTCCATGCTCCACAGGAGGGCAAGAACCAAGGTCCCGGTGCATATGCA 2387

QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGCTTTGGT 2447

QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCCAGGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAAT 2507

QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCACTGGCAGCCCCATCAGTACTCCACTACGGCAAGTTCTTCCGACGGCGGGGTGC 2567

QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGGCGCTTATGACATAATAATTTGTGACGAGTGGCCACTCCACGGATGCCACATCC 2627

QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTTGGGCATCGGCATGCTCTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTTGTG 2687

QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTGGCACCCGCCACCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGGAGGTT 2747

QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2748 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC 2807

QY 361 LysGlyGlyArgHisIleuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGGAGACATCTCATCTTCTGTCAATTCAAAGAGAAGTGCAGCAACTCGCCGCA 2867

QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGGTCCANTTGGGCATCAATGCGTGGCTTACTTACCGCGGTCTTTCAGTGTCCGTC 2927
```

```
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCCGACAGGGCGCATGTGTGTCGTGGCAACGATCCCTCATGACCGCTATACC 2987

QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GCGCACTTCGACTCGGTGATAGACTCAATACGTGTGTGTCCACGACAGTCGATTTTCAGC 3047

QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCCCTTACCTTACCATTGAGACAATCAGCTCCCCCAGGATGCTGTCTCCCGACT 3107

QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 3108 CAACTGCGGGGAGAGACTGGCAGGGGGAAGCAGGCACTCTACAGATTTGTGGCACCGGGG 3167

QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCGGCATGTTTCGACTCGTCCGCTCTGTGAGTGTCTATGACGAGGCTGT 3227

QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 3228 GCTTGTATAGACTACGCCCGCCGAGACTACAGTTAGGCTACGACGGTACATGAACACC 3287

QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCGCTGTGCCAGGACCATCTTGAATTTTGGGAGGGCTCTTTACAGGCCCTC 3347

QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 3407

QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db 3408 CTGGTAGCTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCTCCCCCATCTGCTGGGAC 3467

QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGGAAGTGTGTGATTGCTCCCTCAAGCCACCCCTCCATGGGCCAACACCCCTGCTA 3527

QY 601 TyrArgGlyGlyAlaValalGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3528 TACAGACTCGGGCGCTGTTCAGAAATGAATCAACCTGACGCCACCCAGTACCAAAATACATC 3587

QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3588 ATGACATGATGTTCGGCCGACCTGGAGGTCTGCACGAGCACCTGGGTGCTCGTGGCGGC 3647

QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTGGCTGCTTTGGCGCGTATTGCTGTCAACAGGCTGCGTGTGTCATAGTGGCAGG 3707

QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTCTGCTTGTTCGGGAAGCGGCAATCATACCTGACAGGGAATGCTCTTACCGAGAGTTC 3767

QY 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785

RESULT 9
AAA75296
ID AAA75296 standard; cDNA; 8316 BP.
XX
AC AAA75296;
XX
XX 15-JAN-2001 (first entry)
DT
XX cDNA sequence compiled Hepatitis C virus cDNA clones.
DE
XX Hepatitis C virus; HCV; antisense polynucleotide; polypeptide;
KW viral infectivity; viral replication; ds.
XX
```


Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 3643 GGCTCTCGGACCTTTACCTGGTACAGGACGCGCATGTCTCCCGTGGCGGG 3702
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuValGlySerSer 140
Db 3703 GGTGATACAGGGGAGCGCTGCTGCGCCCGGCCCATTTCTACTTGAAGGCTCCGTCG 3762
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3763 GGGGGTCCGCTGTTGTGCCCCCGGGGACGCGGTGGGCATATTAGGGCCGCGGTGTC 3822
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3823 ACCGCTGGAGTGGCTAAAGCGGTGACATTTATCCCTGTGGAGAACCTAGACACAACCATG 3882
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGluVal 200
Db 3883 AGGTCCCCGGGTGTTTACGGATAACTCTCTCCACGAGTAGTGCCCCAGAGCTTCCAGGTG 3942
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 3943 GCTCACCTCCATGCTCCACAGGCGCGGCAAGACCAAGGTCCCGGCTGCATATGCA 4002
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4003 GCTCAGGCTATAAGGTCTAGTACTCAACCCCTCTGTGTGCAACACTGGGCTTTGGT 4062
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4063 GCTTACATGTCACAGGCTCATGGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAAAT 4122
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyLysCys 280
Db 4123 ACCACTGGCACCCTCATCGTACTTCCACCTACGGCAAGTTCCTTCCGACGGCGGGTGC 4182
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4183 TCGGGGGCGCTTATGACATAATAATTTGTACAGAGTGCCACTCCACGGATGCCACATCC 4242
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4243 ATCTTGGGCACTCGGCACTGCTCTTGACCAAGCAGAGACTGCGGGGGGAGACTCGTTGGT 4302
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4303 CTCCGCCACCGCACCCCTCCGGGCTCCGTACTGTGCCCCATCCCAACATCGAGAGGTT 4362
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4363 GCTCTGTCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTAATC 4422
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4423 AAGGGGGGAGACATCTCATCTCTGTCAATCAAGAGAGAGAGTGCAGCACTCGCGCA 4482
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4483 AAGCTGGTCGATTTGGGATCAATCGCGTGGCCCTACTACCGCGTCTTTCAGTTCGCTC 4542
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4543 ATCCCGACCGCGGATGTTGTGCTGTCGCAACCGATGCTCATCGCGCTATACC 4602
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4603 GGGCACTTCGACTCGGTATAGACTGCAATACGTGTGTACCCAGACAGTTCGATTTTCAGC 4662
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4663 CTTGACCCCTACCTTTACCATTTGAGACAATACGCTCCCGCCAGGATGTGTCTCCCGCACT 4722

Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4723 CAACGTGGGGCAGGACTGGCAGGGGAAGCAGGCATCTACAGATTTGTGGCACCGGG 4782
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4783 GAGCGCCCTCCGCGCATGTTGACTCGTCCGTCTCTGTGAGTGTATGACGCGAGGCTGT 4842
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4843 GCTTGGTATGAGCTACGCGCCGCGGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4902
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 4903 CCGGGGCTTCGCGTGTCCAGGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCCTC 4962
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 4963 ACTCATATAGATGCCACTTTCATCCAGACAAGCAGAGTGGGGAGAACCTTCTCTTAC 5022
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 5023 CTGGTAGGTACCAAGCACCGGTGTGCGCTAGGCTCAAGCCCTCCCCCATCGTGGGAC 5082
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5083 CAGATGTGGAAGTGTGATTCGCTCAAGCCCCACCTCCATGGGCCCAACACCCCTGCTA 5142
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5143 TACAGACTGGGCGCTGTTCAGAAATCACCTGACGACCCAGTCCCAATAACATC 5202
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5203 ATGACATGCAATGTGGCGGACCTGGAGGTGCTCAGCAGACCTGGGTGCTCGTTGGCGGC 5262
Qy 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5263 GTCTGGCTGCTTTGGCGCGTATTGCTGTCAACAGGCTGCGTGGTTCATAGTGGCAGG 5322
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5323 GTGCTCTGTTCGGGGAAGCGGCAATCATCTGACAGGGAAGTCTCTACCGAGAGTTC 5382
Qy 681 AspGluMetGluGluCys 686
Db 5383 GATGAGATGGAAGAGTGC 5400
RESULT 11
AAQ10566
ID AAQ10566 standard; DNA; 9185 BP.
XX AAQ10566;
AC AAQ10566;
XX 25-MAR-2003 (revised)
DT 29-APR-1991 (first entry)
XX Hepatitis C virus strain 1 DNA.
DE Hepatitis C virus;
XX Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;
KW viral infections; ss.
XX Hepatitis C virus.
OS
XX EF414475-A.
PN
XX 27-FEB-1991.
PD
XX
PF 21-AUG-1990; 90EP-00309120.
XX
PR 25-AUG-1989; 89US-00398667.
XX
PA (CHIR) CHIRON CORP.

QY 501 GlnMetTrpLysCysLeuIleAArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 DB 5135 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCACCTCTCCATGGGCCAACACCCCTGCTA 5134
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTrile 620
 DB 5195 TACAGACTGGCGCTGTTGAGATGAAATCATCCTTCAGCGACCCAGTCCACAAATACATC 5254
 QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 DB 5255 ATGACATGCATGTGGCGGACCTGGAGGTGTCACGAGCACCTGGGTCTGTTGGCGGC 5314
 QY 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValLeuValGlyArg 660
 DB 5315 GTCTCGTGTGTTGGCGCGCTATTGCTGTCAACAGGCTGCTGTGTCATAGTGGCAGG 5374
 QY 661 IleValLeuSerGlyLysProAlaIleProAspArgGluValLeuTyrArgGluPhe 680
 DB 5375 GTCTGTCTTGTCCGGGAAGCCGCAATCATCTGACAGGAAGTCTCTACCGAGAGTTC 5434
 QY 681 AspGluMetGluGluCys 686
 DB 5435 GATGAGATGGAGAGTGC 5452

RESULT 12

AAA75297
 ID AAA75297 standard; cDNA; 9185 BP.

XX AC AAA75297;

XX DT 15-JAN-2001 (first entry)

XX DE Sense strand of HCV encoding a polyprotein.

XX KW Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;

XX OS Hepatitis C virus.

XX FH Key Location/Qualifiers

XX FT CDS 320..9184

XX FT /*tag= a

XX FT /note= "partial sequence; no termination codon given"

XX EP1034785-A2.

XX PD 13-SEP-2000.

XX PF 16-MAR-1990; 2000EP-00109602.

XX PR 17-MAR-1989; 89US-00325338.

XX PR 20-APR-1989; 89US-00341334.

XX PR 18-MAY-1989; 89US-00355002.

XX PR 16-MAR-1990; 90EP-00302866.

XX PA (CHIR) CHIRON CORP.

XX PI Houghton M, Choo Q, Kuo G;

XX DR WPI; 2000-566891/53.

XX DR P-PSDB; AAB18541.

XX PT Novel composition comprising a hepatitis C virus antisense polynucleotide

XX PT which is complementary to or corresponds to a sense strand of the virus

XX PT genome, and selectively hybridizes to it.

XX PS Example; Fig 17; 75pp; English.

XX CC The specification describes a pharmaceutical composition which comprises

XX CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is

XX CC characterized by a positive stranded RNA genome which has 40% homology at

XX CC the polypeptide level to a HCV polyprotein. The antisense polynucleotide

CC binds to cellular polynucleotides which enhance and/or are required for
 CC viral infectivity, replicative ability or chronicity. The antisense
 CC polynucleotides may also be designed to bind with high specificity, to be
 CC of increased stability, to be stable and to have low toxicity. The
 CC composition also comprises an agent which causes viral RNA to be
 CC inactive. The composition is used for preventing HCV replication in a
 CC system. The present sequence represents a novel HCV cDNA sequence, which
 CC is used in the course of the invention

XX SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,48e-241 Length: 9185
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x AAA75297 (1-9185)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20

DB 3395 CTGGCGCCCATCAGCGGTACGCCACAGACAGAGGGGCTCTTAGGGTGATATATCACC 3454

QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40

DB 3455 AGCCTAACTGCGCGGACAAAACCAAGTGAGGGTGCAGATTGTGTCAACTGCT 3514

QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60

DB 3515 GCCCAAACTTCTCGGCAACGTGCATCAATGGGGTGTCTACACGGGGGCC 3574

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

DB 3575 GGAACGAGGACCATCGCGTCACCCAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 3634

QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100

DB 3635 CAAGACCTTGTGGCTGGCGGCTCGCAAGTAGCGCTCATTGACACCCCTGCACATTGC 3694

QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120

DB 3695 GGCTCTCTCGGACCTTACCTGGTCACGAGGCACCGCATGTCTATCCGTCGCGCGCGG 3754

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140

DB 3755 GGTGATAGCAGGGGACGCTGTGTGCGCCCGGCCCATTTCTTACTTGAAGGCTCTCTCG 3814

QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160

DB 3815 GGGGTCGCTGTGTGTCGCCCGCGGGGACGCGGTGGGCATATTTAGGGCGCGGTGTGC 3874

QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180

DB 3875 ACCCGTGGAGTGGCTAAGGGGTGACTTTATCTCTGTGGAGAACCTAGACACCAATG 3934

QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200

DB 3935 AGGTCCCGGTGTTCACGGATACTCTCTCCACAGTAGTGGCCCGAGGCTCCAGGTG 3994

QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220

DB 3995 GCTCACCTCCATGCTCCACAGGCGGCGCAAGACCAAGGTCCTGGCTGCATATGCA 4054

QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240

DB 4055 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTTGGT 4114

QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260

DB 4115 GCTTACATGTCCAGGCTCATGGATCGATCCTTAACTACAGGACCGGGGTGAGAACAT 4174

```
QY 261 ThrThrGlySerProIleThrThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 4175 ACCACTGGCAGCCCATCACTACCTCCACCTACGGCAAGTTCCTTGCAGACGGCGGTGC 4234
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 4235 TCGGGGGGCGCTTATGACATAAATAATTTGTGACAGTGGCCACTCCACGGATGCCACATCC 4294
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 4295 ATCTTGGGCATCGGCACTGCTTGTACCAAGCAGAGACTCGGGGGGAGACTGGTTGTG 4354
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 4355 CTCGCCACCGCCACCCCTCGGGCTCGCTCACTGTGCCCATCCCAACATCGAGGAGTT 4414
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 4415 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC 4474
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
DB 4475 AAGGGGGGAGACATCTCACTCTCTGTCAATCAAGAGAGTGGCAGCACTCGCGCA 4534
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 4535 AAGCTGTGTCATTTGGCATCAATCGCGTCCGCTACTACCGGGTCTTGACGTGTCCGTC 4594
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 4595 ATCCCGACCAACGGCGGATGTTGTGTGTGGCAACCGATGCCCTCATGACCGGCTATACC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 4655 GGGCACTTCGACTCGGTAGACTGCATACGTGTGTACCCAGACAGTCGATTTTCAGC 4714
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 4715 CTTGACCTTACCTTCCACTTGAGACAATCACGCTCCCCCAGGATGCTGTCTCCGCACT 4774
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 4775 CAACGTGGGGGAGAGCTGGCAGGGGAGCCAGGCACTTACAGATTTTGTGGCACCGGG 4834
QY 481 GluArgProSerGlyMetPheAspSerValIleCysGluCysTyrAspAlaGlyCys 500
DB 4835 GAGCGCCCTCCGCGATGTCGACTGCTCGTCTCTGTGAGTGTATGACGAGGCTGT 4894
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 4895 GCITGGTATGAGCTCAGCCCGCGGAGACTACAGTTAGGCTACGAGCTACATGAACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
DB 4955 CCGGGGCTCCCGTGTGCCAGGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGCGCTC 5014
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 5015 ACTCATATAGATGCCACTTTCATCCAGAAAGCAGAGTGGGAGAACCTTCTCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
DB 5075 CTGGTAGGTACCAAGCACCGGTGCGCTAGGCTCAAGCCCTCCCCCATCGTGGAC 5134
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 5135 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCCAACCCCTCCATGGGCCAACACCCCTGCTA 5194
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrThrHisProValThrLysTyrIle 620
DB 5195 TACAGACTGGGGCGTGTTCAGATGAATGAATCACCCTGACGACCACCATCAATATCATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
```

```
DB 5255 ATGACATGATGTGCGCCACCTGGAGGTGCTCACGAGACCTGGGTGCTGTTGGCGC 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValIleValGlyArg 660
DB 5315 GTCCTGGCTGCTTTGGCCGCTATTGCCCTGTCAACAGGCTGCGTGTCTATAGTGGCAGG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 5375 GTCGCTTCTCGGGAAGCCGCAATCATACCTGACAGGAAGTCTCTTACCGAGAGTTC 5434
QY 681 AspGluMetGluGluCys 686
DB 5435 GATGAGATGAAGAGTGC 5452
RESULT 13
ADN35979
ID ADN35979 standard; cDNA; 9185 BP.
XX
AC ADN35979;
DT 17-JUN-2004 (first entry)
XX
DE HCV cDNA clone #2.
XX
KW Antiviral; Vaccine; hepatitis C virus infection; HCV infection; ss.
XX
OS Hepatitis C virus.
XX
PN EP1394255-A2.
XX
PD 03-MAR-2004.
XX
PF 16-MAR-1990; 2003EP-00016585.
XX
PR 17-MAR-1989; 89US-00325338.
PR 20-APR-1989; 89US-00341334.
PR 18-MAY-1989; 89US-00355002.
PR 16-MAR-1990; 90EP-00302866.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Houghton M, Choo Q, Kuo G;
XX
WPI: 2004-193149/19.
DR P-PSDB; ADN35978.
XX
PT Novel purified hepatitis C virus polypeptide comprising epitope encoded
PT by hepatitis C virus cDNA, useful as vaccine for treating hepatitis C
virus.
XX
PS Claim 1; Fig 17; 79pp; English.
XX
CC The present invention relates to hepatitis C virus (HCV) proteins and
CC cDNA sequences. The sequences are useful in immunoassays for detecting
CC antibodies directed against HCV antigen; preparing host cells transformed
CC with a recombinant polynucleotide; screening antiviral agents and
CC determining the effect of antiviral agent in inhibiting viral replication
CC in cell culture system; and developing vaccine for treating HCV
CC infection.
XX
SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,48e-241 Length: 9185
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 12 Gaps: 0
US-09-930-591-2 (1-686) x ADN35979 (1-9185)
```

```
Qy 1 MetAlaProIleThrAlaTyAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db CTGGCGCCCATCAGCGGTAGCCAGCAGACAAGGGCCCTCTAGGGTGCAATACCC 3454
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db AGCCTAAGTGGCCGGGACAAACCAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3514
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyHisGlyAla 60
Db GCCCAAACTTCCTGGCAACGTGCATCAATGGGGTGTGCTGACTGTCTACCAACGGGGCC 3574
Qy 61 GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetTyThrAsnValAsp 80
Db GGAACGAGGACCATCGGCTCACCAAGGGTCTCTGTATCCAGATGTATACCAATGTAGAC 3634
Qy 81 GlnAspLeuValGlyTTPProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db CAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCGCTCATTTGACACCCCTGCACATTGC 3694
Qy 101 GlySerSerAspLeuTyLeuValThrArgHisAlaAspValIleProValArgArg 120
Db GGCTCCTCGGACCTTACCTGGTCACGAGGCACGCCGATGTTCATTCCTGGCGCGGG 3754
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyLeuLysGlySer 140
Db GGTGATAGCAGGGCAGCTGCTGTCGCCCGGCCATTTCTACTTTGAAAGGCTCCTCG 3814
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db GGGGGTCCGCTGTGTGCCCCCGGGGCACGCCGTGGGCATATTAGGGCGCGGTGTC 3874
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db ACCCGTGGAGTGGCTAGGCGGTGGATTTATCTCTGTGGAGAACCTTAGACAAACCATG 3934
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyGlnVal 200
Db AGGTCCCGGGTGTTCAGGATAACTCTCTCCACCAGTAGTGTGCCCCAGAGCTCCAGGTG 3994
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyAla 220
Db GCTCACCTCCATGTCTCCACAGGACGGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA 4054
Qy 221 AlaGlnGlyTyLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db GCTCAGGGCTATAGGTGTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGGCTTTGGT 4114
Qy 241 AlaTyMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db GCTTACATGTCGAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 4174
Qy 261 ThrThrGlySerProIleThrTySerThrTyGlyLysPheLeuAlaAspGlyGlyCys 280
Db ACCATGGCAGCCCATCATCGTACTCCACCTACGGCAAGTTCTTTCGCCAGCGGGGTGC 4234
Qy 281 SerGlyGlyAlaTyAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db TCGGGGGCGGCTTATGACATAATAATTGTGACAGGTGCCACTCCAGGATGCCACATCC 4294
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db ATCTTGGGCATCGGCACTGCTTGTACCAAGCAGAGACTGGGGGGCGAGACTGGTTGTG 4354
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db CTCGCCACCGCACCCCTCCGGGCTCGGTCACTGTGCCCTCACTCCCAACATCGAGGAGTT 4414
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyGlyLysAlaIleProLeuGluAlaIle 360
Db GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAACT 4474
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
```

RESULT 14

AAT12710

ID AAT12710 standard; cDNA; 9401 BP.

XX

AC AAT12710;

XX

```
Db 4475 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAAGAGTGCAGAACTCCGCCA 4534
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyTrpArgGlyLeuAspValSerVal 400
Db 4535 AGCTGTGTGCATTTGGGCATCAATGCCGTGCCCTACTACCGCGTCTTGAGCTGTCCGTC 4594
Qy 401 IleProThrSerGlyAspValValValAlaLathrAspAlaLeuMetThrGlyPheThr 420
Db 4595 ATCCCGACCAAGCGCGATGTTGTCGTGTGGCAACCGATGCCCTCATGACCGGCTATACC 4654
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4655 GGGCACTTCGACTCGGTGTAGACTGCAATACGTGTGTACCCACAGACAGTGCATTTCCAGC 4714
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTTGACCTTACCTTCCACATTTAGACAATCAGCTCCCCCAGGATGCTGTCTCCCGCACT 4774
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyArgPheValAlaProGly 480
Db 4775 CAACGTCGGGGCAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGG 4834
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyAspAlaGlyCys 500
Db 4835 GAGCGCCCTCCGGCATGTTTCGACTCGTCCGTCTCTGTGAGTGTCTATGACGCGAGGCTGT 4894
Qy 501 AlaTrpTyTrpGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyMetAsnThr 520
Db 4895 GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTTACGAGCGTATCATGAAACAC 4954
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGGCTTCCGCTGTGCCAGGACCATCTTGAATTTTGGGAGGGCGTCTTTTACAGGCTC 5014
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTy 560
Db 5015 ACTCATATAGATGCCCACTTCTTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTCTTAC 5074
Qy 561 LeuValAlaTyTrpGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 5075 CTGGTAGCTTACCAAGCACCGGTGTGGCTTAGGGCTCAAGCCCTCCCATCGTGGGAC 5134
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGGAGTGTGTGATTCCGCTCAAGCCCAACCTCCATGGGCCCAACACCTCTGCTA 5194
Qy 601 TyArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyIle 620
Db 5195 TACAGACTGGGCGCTGTTTCAAGATGAATCAACCTGACGCACCCAGTCACCAATAATCATC 5254
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5255 ATGACATGCAATGTGGCCGACCTGGAGGTGCTCAGCAGCACCTGGGTGCTCGTTGGCGGC 5314
Qy 641 ValLeuAlaAlaLeuAlaTyCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTGGTCTGTTTGGCCGCGTATTGCTGTCAACAGGCTGCGTGTCTATAGTGGGAGG 5374
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyArgGluPhe 680
Db 5375 GTGCTCTTGTCCGGGAAGCGCGCAATCATACCTGACAGGGAAGTCTCTCTACCGAGAGTTC 5434
Qy 681 AspGluMetGluCys 686
Db 5435 GATGAGATGGAAGAGTGC 5452
```

RESULT 14

AAT12710

ID AAT12710 standard; cDNA; 9401 BP.

XX

AC AAT12710;

XX

DT 25-MAR-2003 (revised)
 XX 15-MAY-1996 (first entry)
 DE Hepatitis C virus polyprotein.
 XX Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;
 KW antibodies; ds.
 KW Hepatitis C virus.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 342..9378
 FT /*tag= a
 XX
 XX EP693687-A1.
 XX
 XX 24-JAN-1996.
 XX
 XX 03-APR-1991; 95EP-00114016.
 XX
 XX 04-APR-1990; 90US-00504352.
 XX
 XX (CHIR) CHIRON CORP.
 XX
 XX Houghton M, Choo Q, Kuo G;
 XX PI
 XX WPI; 1996-117956/13.
 XX P-PSDB; AAR90931.
 XX
 XX Combinations of synthetic Hepatitis C Virus antigens - provide more
 PT effective diagnosis of Non-A, Non-B Hepatitis.
 PT
 XX Disclosure; Fig 1(A-Y); 53pp; English.
 XX
 CC The combination comprises an HCV antigen from the C domain (pref. C22 -
 CC AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -
 CC AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5
 CC (AAR90934) domain. The antigens may in the form of a fusion protein, a
 CC simple physical mixture, or the individual antigens commonly bound to a
 CC solid matrix. They are pref. prep'd by recombinant DNA techniques
 CC (primers are given in AAR12711-112716), but can be synthesised or
 CC isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003
 CC to correct PF field.)
 XX
 SQ Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,55e-241 Length: 9401
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 2 Gaps: 0

US-09-930-591-2 (1-686) x AAT12710 (1-9401)
 QY 1 MetAlaProIleThrAlaTrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 3417 CTGGCGCCCATCAGCGGTACGCCACAGACAGAGGGGCTCTTAGGGTGCAATCACC 3476
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 DB 3477 AGCCTAACTGGCGGGACAAAACCAAGTGGAGGGTGAGTCCAGATTGTGCAACTGCT 3536
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTrHisGlyAla 60
 DB 3537 GCCCAAAACCTTCTGGCAACGTCATCAATGGGGTGCTGGGACTGTCTACCACGGGGCC 3596
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTrpThrAsnValAsp 80
 DB 3597 GGNACAGAGNACCATCGGTGTCACCAAGGTCCTGTATCATCAAGATGATACCAATGAGAC 3656
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100

DB 3657 CAAGACCTTGTGGCTGGCCCGCTCGCAAGTAGCGGCTCATTTGACACCCCTGCACCTGC 3716
 QY 101 GlySerSerAspLeuTyrIleuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 3717 GGCTCTCGGACCTTTACCTGGTCAGAGCAGCGCATGTCTATTCCTCGCGCGCGCG 3776
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 DB 3777 GGTGATAGCAGGGGAGCCTGCTGCGCCCGGCCCATTTCTTCTTGAAGGCTCCTCG 3836
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 3837 GGGGGTCCGCTGTGTGCCCCCGGGGACGCGCTGGGCATATTTAGGCCCGCGGTGTGC 3896
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 DB 3897 ACCCGTGGAGTGGCTAAGCGGTGACCTTTATCCCTGTGTGAGAACCTGAGACCAACATG 3956
 QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
 DB 3957 AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTCCCGCCAGAGCTTCCAGGTG 4016
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 4017 GCTCACTCCATGCTCCACAGCAGCGGCAAAAGCACCAGGTCCTCGGCTGCATATGCA 4076
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 4077 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACATCGGGCTTGTG 4136
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 4137 GCTTACATGTCAGGCTCATGGATCGATCTCTACATCAGGACCGGGGTGAGAACAT 4196
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 4197 ACCACTGGCAGCCCATCAGTACTCCACTACGCCAAGTTCCTTGGCCGACGCGGGTGC 4256
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGlyCysHisSerThrAspIleThrSer 300
 DB 4257 TCGGGGGGCGCTTATGACATATATTTGTGACGAGTGCCACTCCACGGATGCCACATCC 4316
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 4317 ATCTTGGGCATCGGCATGCTCTTACCCACAGAGAGACTGCGGGGCGAGACTGGTTGTG 4376
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 4377 CTCGCCACCGCCACCCCTCGGGCTCCGTCACCTGTGCCCATCCCAACATCGAGAGGTT 4436
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 DB 4437 GCTCTGTCCACCACCGAGAGATCCCTTTTACGCCAAGGCTATCCCTCCGAAGTAATC 4496
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
 DB 4497 AAGGGGGGGAGACATCTATCTTCTGTCATCAAGAGAGAGTGCAGCAACTCCCGCA 4556
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 4557 AAGCTGGTCGATTGGGCATCAATCCGTGGCTACTACCGCGGTCTTGACGTGCGGTC 4616
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 4617 ATCCCGACAGCGGGCATGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGCTATACC 4676
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 4677 GCGGACTTCGACTCGGTGATAGACTGCAATAGCTGTGTCCACAGACAGTGCATTTCCAG 4736
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460


```

Db      4737 CTTGACCCCTACCTTACCATTGAGACAAATCAGCTCCCGGAGTGTCTCCCGCACT 4796
Qy      461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db      4797 CAACGTGGGGCAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCAGGG 4856
Qy      481 GluArgProSerGlyMetPheAspSerSerValLeuGlyGluCysTyrAspAlaGlyCys 500
Db      4857 GAGCGCCCTCCGGCATGTTGACTCGCTCCGCTCTGTGAGTGTATGACGCGAGGCTGT 4916
Qy      501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db      4917 GCTTGATAGTACTCACCCGCCGAGACTACAGTTAGCTACGAGCGTACATGAACACC 4976
Qy      521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db      4977 CGGGGCTTCCGTTGTCAGGACCATCTTGAATTTTGGAGGGGCTTTTACAGGCCTC 5036
Qy      541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db      5037 ACTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTCTTAC 5096
Qy      561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db      5097 CTGGTAGCGTACCAGCCACCGTGTGGCTAGGCGCTCAAGCCCTTCCCCCATCGTGGGAC 5156
Qy      581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      5157 CAGATGTGGAGTGTGTTGATTCGCTCAAGCCCACTTCCATGGCCCAACACCTTGCTA 5216
Qy      601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      5217 TACAGACTGGCGCTGTTTCAGAAATCAGAAATCAGCTGAGCCACCCAGTCACCAATACATC 5276
Qy      621 MetThrCysMetSerAlaAspLeuValValThrSerThrTyrValLeuValGlyGly 640
Db      5277 ATGACATGATGTCGGCGAGCTGGAGTGTGTCAGAGCACCTGGGTCTCGTTGGCGGC 5336
Qy      641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      5337 GTCTGTGCTGTTGGCGCGTATGCTGTCACAGGCTGCTGGTGTATAGTGGGAGG 5396
Qy      661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db      5397 GTCGTGTTGTCGGGAGCGGCAATCATCTGACAGGAAGTCTCTACCGAGAGTTC 5456
Qy      681 AspGluMetGluGluCys 686
Db      5457 GATGAGTGGAGAGTGC 5474
RESULT 15
AAT99981
ID      AAT99981 standard; DNA; 9401 BP.
XX
AC      AAT99981;
XX
DT      25-MAR-2003 (revised)
DT      16-MAR-1998 (first entry)
XX
DE      HCV polyprotein coding sequence.
XX
KW      PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
KW      C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
KW      NS4; ds.
XX
OS      Hepatitis C virus.
XX
FH      .Key      Location/Qualifiers
FT      CDS      342..9377
FT      /*tag= a
XX
PN      US5683864-A.
XX

```

04-NOV-1997.

07-JUL-1992; 92US-00910760.

18-NOV-1987; 87US-00122714.

30-DEC-1987; 87US-00139886.

26-FEB-1988; 88US-00161072.

06-MAY-1988; 88US-00191263.

26-OCT-1988; 88US-00263584.

14-NOV-1988; 88US-00271450.

17-MAR-1989; 89US-00325338.

20-APR-1989; 89US-00341334.

21-APR-1989; 89US-00353896.

18-MAY-1989; 89US-00355002.

04-APR-1990; 90US-00504352.

(CHIR) CHIRON CORP.

Kuo G, Houghton M, Choo Q;

WPI; 1997-548976/50.

P-PSDB; AAW34480.

Combination of three hepatitis C virus antigens - used for detection of specific antibodies to diagnose infection.

Disclosure; Col 25-46; 57pp; English.

This sequence represents the Hepatitis C virus polyprotein coding sequence. Fragments of this sequence can be amplified and used in the combination of HCV antigens of the invention. The HCV antigen combination comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa) 1-120 of the HCV polyprotein), or its immunologically reactive fragment containing at least 8 aa. It also comprises two additional antigens from two different polyprotein domains, including at least 8 aa from the NS3, NS4, S or NS5 domains of the polyprotein, corresponding, respectively, to aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polyprotein. Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa regions of the HCV polyprotein. These antigen combinations are used diagnostically to detect anti-HCV antibodies, using any standard immunoassay format. These antigen combinations have a broader range of reactivity with antibodies than any antigen individually. (Updated on 25-MAR-2003 to correct PR field.)

Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.:	2.55e-241	Length:	9401
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	2	Gaps:	0

US-09-930-591-2 (1-686) x AAT99981 (1-9401)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20

Db 3417 CTGGCGCCCATCAGCGGTACGCCAGACAGAGGGCCCTCTAGGTGTCATATACCC 3476

Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40

Db 3477 AGCTTAATCTGGCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 3536

Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60

Db 3537 GCCCAAAACCTTCTTGGCAACGTGCATCAATGGGGTGTGTGCTACTGTCTACACGGGCC 3596

Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

Db 3597 GGAACGAGGACCATCGCGTCAACCAAGGGCTCTCATCCAGATGTATACCAATGTAGAC 3656

Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100

```
Db 3657 CAAGACCTTGTGGCGCTGGCCGCTCCGCAAGGTAGCGCTCATTTACACCCCTGCACTTGC 3716
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 3717 GGTCTCTCGGACCTTTATACCTGGTTCACAGGACGCGCATGTCATTCCCGTGCGCCGCGG 3776
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuTysGlySerSer 140
Db 3777 GGTGATAGAGGGGACGCTGCTGTGCGCCCGGCCCATTTCTTACTTTGAAGAGCTCTCTG 3836
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3837 GGGGCTCCGCTGTGTGCTCCCGCGGGCACCGCTGGGCATATTTAGGGCGCGCGGTGTC 3896
Qy 161 ThrArgGlyValAlaIlysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3897 ACCCGTGGAGTGGCTAAGCGGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 3956
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3957 AGTCCCCCGTGTTCACGGATTAACCTCTCTCCACAGTAGTGCCCGCAGAGCTTCCAGGTG 4016
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIaIaTyrAla 220
Db 4017 GCTCACCTCCATGCTCCACAGCGCGGCAAGCACCAAGTCCCGGCTGCATATGCA 4076
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaIaIaThrMetGlyPheGly 240
Db 4077 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGCTTTGGT 4136
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4137 GCTTACATGTCCAAGGCTCATGGATGCATCTTAACATCAGGACCGGGGTGAGAACATT 4196
Qy 261 ThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4197 ACCACTGGCAGCCCCATCAGTACTCCACCTACGGCAAGTTCTTGGCGAGCGGGGTGC 4256
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4257 TCGGGGGGCGCTTATGACATTAATAATTTGTGACAGTGGCCACTCCACGGATGCCACATCC 4316
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4317 ATCTTGGGCATCGGCACTGCTTGTGACCAAGAGAGACTCGGGGGGCGAGACTGGTTGTG 4376
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4377 CTGCCACCGCCACCCCTCCGGGCTCCGTCCTCACTGTGCCCATCCCAACATCGAGGAGTT 4436
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4437 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTAATC 4496
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 4497 AAGGGGGGAGACATCTCATCTCTGTCTATTCAAGNAGAGTCCGACGAACTCGCGCA 4556
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4557 AAGCTGTGTGCATTGGGCATCAATCGCGTCCGTCCTACTACCGCGGTCTTTGACGTGTCGTC 4616
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4617 ATCCCGACCAAGCGCGATGTGTGCTGTGCGCAACCGATGCCCTCATGACCGGTATACC 4676
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4677 GGGGACTTTCGACTCGGTGATAGTGCATATGATGCTGTGTACCCAGACAGTTCGATTCAGC 4736
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
```

```
Db 4737 CTTGACCCCTACCTTACCATTGAGCAATACAGCTCCCCCAGGATGCTGTCTCCCGCACT 4796
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4797 CAACGTGGGGGAGGACTGGGAGGGAAGCCAGGCACTTACAGATTGTGGACCCGGGG 4856
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4857 GAGGCGCCCTCCGCGCATGTTCCGACTCGTCCGCTCTGTGAGTGTATGACGCGAGCTGT 4916
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4917 GCTTGTGTATGAGTCAACGCGCCGCGAGACTACAGTTAGGCTACGAGCTACATGAACAC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 4977 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGCTTTTACAGGCTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5037 ACTCATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerThrAsp 580
Db 5097 CTGGTAGCTTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCCCTCCCCCATCGTGGAC 5156
Qy 581 GlnMetTyrLysCysIleuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5157 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCCCACCTCCATGGGCCAACACCCCTTGCTA 5216
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5217 TACAGACTGGGCGCTGTTCAGAATGAATCAACCTGACGACCCAGTCACCAATACATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 5277 ATGACATGCAATGTCCGCCACCTGGAGGTGCTCACGAGCACCTGGGTGCTCGTTGGCGG 5336
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5337 GTCTTGGCTGCTTTGGCCCGGTATTCCTTCAACAGGCTGCTGTGCTATAGTGGCAGG 5396
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTCGTCTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGATGGAAGAGTGC 5474
```

Search completed: February 27, 2005, 04:08:03
Job time : 946 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2005, 03:44:53 ; Search time 296 Seconds
(without alignments)
3792.182 Million cell updates/sec

Title: US-09-930-591-2

Perfect score: 3618

Sequence: 1 MAPITAAQTRGLGCIIT.....PALIPDRELYREFDEMEEC 686

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/uspto.spool/p/US0930591/runat.25022005.143613.25700/app.query.fasta.1.839
-DB=Issued Patents NA -OPMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCT=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0930591.qcgn.1.105 @runat.25022005.143613.25700 -NCPUS=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXI=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3574	98.8	7310	3	US-08-444-818-74
2	3574	98.8	9379	3	US-09-388-874-1
3	3574	98.8	9379	4	US-09-916-359-1
4	3574	98.8	9401	1	US-07-910-760-9
5	3574	98.8	9401	1	US-08-440-519-9
6	3574	98.8	9401	3	US-08-440-549-9
7	3574	98.8	9401	3	US-08-823-895A-25
8	3571	98.7	6785	3	US-08-444-818-65
9	3570	98.7	8316	3	US-08-444-818-88
10	3570	98.7	9185	3	US-08-444-818-122
c 11	3570	98.7	9185	3	US-08-444-818-123
12	3565	98.5	2058	4	US-09-881-239-2

13	3565	98.5	2058	4	US-09-881-654-1
14	3565	98.5	2058	4	US-10-637-323-1
15	3565	98.5	8987	3	US-08-444-818-137
16	3555	98.3	5360	3	US-08-444-818-53
17	3551	98.1	9646	3	US-08-811-566-1
18	3551	98.1	9646	3	US-09-034-756-1
19	3551	98.1	12980	3	US-08-811-566-5
20	3551	98.1	12980	3	US-09-034-756-5
21	3544	98.0	9599	3	US-09-014-416-2
22	3544	98.0	9599	3	US-09-014-416-6
23	3538	97.8	9401	2	US-08-432-693-1
24	3538	97.8	9416	3	US-08-811-566-19
25	3538	97.8	9416	3	US-08-034-756-19
26	3535	97.7	9379	3	US-08-444-818-176
27	3533	97.7	9365	4	US-09-827-688-7
28	3532	97.6	9401	5	PCT-US91-02225-9
29	3504	96.8	9416	3	US-08-823-895A-26
30	3504	96.8	9416	4	US-10-104-866-13
31	3429	94.8	7989	4	US-09-539-601-10
32	3429	94.8	8001	4	US-09-539-601-7
33	3426	94.7	8637	4	US-09-539-601-4
34	3426	94.7	8638	4	US-10-029-907-7
35	3426	94.7	8638	4	US-10-029-907-24
36	3426	94.7	8638	4	US-10-029-907-25
37	3426	94.7	8639	4	US-10-029-907-1
38	3426	94.7	8643	4	US-10-029-907-4
39	3426	94.7	8649	4	US-09-539-601-13
40	3426	94.7	11076	4	US-09-539-601-1
41	3423	94.6	8001	4	US-09-539-601-22
42	3421	94.6	8638	4	US-10-029-907-6
43	3421	94.6	9413	1	US-09-827-688-6
44	3420	94.5	6039	1	US-08-324-977-11
45	3420	94.5	6039	2	US-08-384-616-11

ALIGNMENTS

RESULT 1
US-08-444-818-74
; Sequence 74, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/444,818
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 74:


```
Db 3588 ATGACATGTCATGTCGGCGACCTGAGGTCGTCTACGACGACCTGGGTCGTGGCGGC 3647
Qy 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValcIlyArg 660
Db 3648 GTCTGCTGCTTTGGCGCGTATGCTGTCAACAGGCTGCTGGTCATAGTGGCAGG 3707
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTCTGCTTTGTCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 3767
Qy 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785

RESULT 2
US-09-388-874-1
; Sequence 1, Application US/09388874
; Patent No. 6284249
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; TITLE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: PWC97-03A
; CURRENT APPLICATION NUMBER: US/09/388,874
; CURRENT FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: PCT/FR98/00448
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 97/02,887
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320)...(9352)
US-09-388-874-1

Alignment Scores:
Pred. No.: 0 Length: 9379
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-09-388-874-1 (1-9379)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3395 CTGGCGCCCATCAGCGGCTAGCCGACGACGACAAAGGGGCTCTAGGGTGATTAATCAC 3454
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerAla 40
Db 3455 AGCTTAAGTGGCGGGGACAAACCAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3514
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysIleThrThrValTyrHisGlyAla 60
Db 3515 GCCCAAACTTCTCTGGCAAGTGCATCAATGGGGTGTGCTGACTGTCTACACCGGGGCC 3574
Qy 61 GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3575 GGAACGAGGACCATCGCTCCACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 3634
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3635 CAAGACCTTGTGGGTGGCCGCTCCGCAAGGTAGCCGCTCATTCACACACCTGCACTTGC 3694
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
```

```
Db 3695 GGCTCTCTCGGACCTTTACCTGGTGTACGAGGACGCGCATGTCTATCCCGTGGCGCGGG 3754
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3755 GGTGATAGCAGGGGAGCCTGCTGTGCCCCCGCCCATTTCTTACTTTGAAAGGCTCTCTCG 3814
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3815 GGGGGTCTGCTGTGTGCCCCCGGGGACGCGGTGGGCATATTTAGGGCGCGCGGTGTC 3874
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3875 ACCCGTGGAGTGGCTAAGCGGTGAGCTTTATCTCTGTGGAGAACCTTAGAGACAACCATG 3934
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3935 AGGTCCCCGGGTTCACGGATAACTCTCTCCACAGTAGTGCCCCAGAGCTTCCAGGTG 3994
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIleTyrAla 220
Db 3995 GCTCACCTCCATGCTCCACAGCGCAGCGGCAAAAGCACCAGGTCCCGGCTGCATATGCA 4054
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaIleThrMetGlyPheGly 240
Db 4055 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACACCTGGGT 4114
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4115 GCTTACATGTCCAAGGCTCATGGATGATCTTAACATCAGGACCGGGGTGAGAACAT 4174
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4175 ACCACTGGCAGCCCATCAGTACTCCACCTACGGCAAGTTCTTGGCGGCGCGGTGTC 4234
Qy 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGGCGCTTATGACATAATAATTTGTGACAGTGCACATCCAGGATGCCACATCC 4294
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTGGGCTATCGGCACTGCTTGTACCAAGCAGAGACTGCGGGGGCGAGACTGGTGTG 4354
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4355 CTGCGCACCGCCACCCCTCCGGGCTCGTCTACTGTGCCCCCATCCCAACATCGAGGAGTT 4414
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4415 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAAATC 4474
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4475 AAGGGGGGAGACATCTCATCTCTGTCTTCAAGAAGAGTGCAGAGAACTCGCGCA 4534
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4535 AAGCTGTGTGCAATTTGGSCATCAATCGCTGCGCTACTACCGCGTCTTTGACGTCTCGTC 4594
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4595 ATCCCGACCAAGCGCATGTTGTGTCGTGGCAACCGATGCCCTCATGACCGGTATACC 4654
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4655 GGGGACTTCCACTCGGTGATAGACTGCATACGTGTGTACCCAGACAGTGCATTTGAGC 4714
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTTGACCTTACCTTCACTTGAACAATCAGCAATCAGCTCCCCCAGGATGCTGTCTCCGCACT 4774
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4775 CAACGTCCGGGCGAGGACTGGCAGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGG 4834
```

Qy	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	4835	GAGCGCCCTCCGGCATGTTTCGACTCGTCGCTCTGTGAGTGCTATGACGAGGCTGT	4894
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	4995	GCTTGGTATGAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACCC	4954
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
Db	4955	CCGGGGCTTCCCGTGTGTCGACGAGCATCTTGAATTTTGGGAGGCGCTCTTTACAGGCCTC	5014
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	5015	ACTCATATAGATGCCACTTCTTATCCAGACAAAGCAGAGTGGGAGAACCTTCCTTAC	5074
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaAargAlaGlnAlaProProProSerTrpAsp	580
Db	5075	CTGCTAGCGTACCAAGCACCGTGTGGCTAGGGCTCAAGCCCCTCCCATCGTGGGAC	5134
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	5135	CAGATGTGGAGTGTGTTGATTCCGCTCAAGCCACCTTCATGGGCCAACACCCCTGCTA	5194
Qy	601	TyrArgLeuGlyAlaValAlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	5195	TACAGACTGGCGCTGTTTCAGATGAATCACCCTGACGACCCAGTCACCAATATACATC	5254
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	5255	ATGACATGCATGTCCGGCGACTCGAGGTCGTACAGAGCACCTGGGTGCTGCTGTGGCGGC	5314
Qy	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	5315	GTCCTGGCTGCTTGGCGCGGTATTGCTGTCAACAGGCTCGGTGTCATAGTGGGCAGG	5374
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	5375	GTCCTCTGTGTCGGGAGCCGGCNAATCATACCTGACAGGAAGTCTCTACCGAGAGTTC	5434
Qy	681	AspGluMetGluGluCys	696
Db	5435	GATCAGATGGAAGAGTGC	5452

RESULT 3

```

US-09-916-359-1
; Sequence No., Application US/09916359
; Patent No. 6538123
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; TITLE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: EPMC97-03A
; CURRENT APPLICATION NUMBER: US/09/916,359
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 09/388,874
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: 97/02,887
; PRIOR FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320) ... (9352)
US-09-916-359-1

```

Alignment Scores:

Alignment scores:	0	Length:	9379
Pred. No.:	0		

Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	4	Gaps:	0
US-09-930-591-2 (1-686) x US-09-916-359-1 (1-9379)			
Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	3395	CTGGCGCCATACAGGGGTHAGCCAGCAGCAAGGGGCTCTTAGGGTCATATCAATCACC	3454
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValIgluGlycluValGlnIleValSerThrAla	40
Db	3455	AGCCTAACTGGCGGGAGAAAAACCAAGTGGAGGGTGAGGTCAGATTGTGTCAACTGCT	3514
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla	60
Db	3515	GCACAAACCTTCTGGCAACGTGCATCAATGGGGGTGTCTGGACTGTCTACCAACGGGGCC	3574
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	3575	GGAAAGAGGACCATCGCGTCACCAAGGGTCTCTATCCAGATGATATACCAATGTAGAC	3634
Qy	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	3635	CAAGACCTTGTGGCTGGCCCGCTCCGCCAAGGTAGCGCTCATTTGACACCTGCACCTGC	3694
Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg	120
Db	3695	GGCTCTCTCGACCTTTACCTGGTTCAGAGGACGCGCGATGTCTATCCGTGCGCCGGCGG	3754
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
Db	3755	GGTGATAGAGGGGACCTGCTGTGCCCGGCCCATTTCTTACTTGAAGGCTCTCTCG	3814
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	3815	GGGGGTCCGCTGTTGTGCCCCCGGGGCACGCGTGGGCATATTTAGGGCCGCGGTGTC	3874
Qy	161	ThrArgGlyValAlaIalysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
Db	3875	ACCCTGTGGAGTGCTTAAGCGGTGGACTTTTATCCCTGTGGAGAACCTTAGAGACAACATG	3934
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
Db	3935	AGGTCCCGGTGTTCAGGATAACTCTCTCCACCAAGTAGTGCCTCCAGAGCTTCCAGGTG	3994
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIalTyrAla	220
Db	3995	GCTCACCTCATGCTCCACAGGCGAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA	4054
Qy	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	4055	GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGCTGCTCAACACACTGGGCTTGTG	4114
Qy	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	4115	GCATTATGTCFAAGGCTATGGGATCGATCTCTAAATCAGGACCGGGGTGAGAACAATT	4174
Qy	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
Db	4175	ACCACTGGGAGCCCCATCAGTACTTCACTACGGCAAGTTCCTTCCCGACGCGGGGTGC	4234
Qy	281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	4235	TCGGGGGGCGCTTATGACATAATAATTGTGACAGTGCCTCCACCGATGCCACATCC	4294
Qy	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	4295	ATCTTGGGCATCGGCATGTCTTGTGACCAAGCAGACACTCGCGGGCGAGACTGGTTGTG	4354
Qy	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340

Db 4355 CTCCGACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCTCCCAACATCGAGAGGTT 4414
Qy 341 AlaleuSerThrThrGlyGluProPheThrGlyLysAlaIleProLeuGluAlaIle 360
Db 4415 GCTCTGTCCACCCAGGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC 4474
Qy 361 LysGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4475 AAGGGGGGAGACATCTCATCTCTGTCAATCAAAAGAGAAAGTGCAGCACTCGCGCA 4534
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaValAlaValAlaValAlaVal 400
Db 4535 AAGCTGTGCGATTTGGGCATCAATGCGGTGCTTACCGCGTCTTGAGTGTCCGTC 4594
Qy 401 IleProThrSerGlyAspValValValAlaValAlaValAlaValAlaValAla 420
Db 4595 ATCCCGACCGCGGATGTTGCTGCTGCAACCGATGCTCATGACCGGTATACC 4654
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4655 GCGCACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTGCATTCAGC 4714
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTTGACCCCTACCTTCACCATTTGACAAATCACTGCTCCCGAGGATGCTCTCCCGCACT 4774
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4775 CNAAGCTGGGCGAGAGCTGGCAGGGGAGCAGGCACTACAGATTTGTGGCACCGGG 4834
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4835 GAGCGCCCTCCGCGATGTTCCGATCGCTCGCTCTGTGAGTGTATGACGCGAGGCTGT 4894
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4895 GCTTGGTATGAGTCACGCGCCGCGAGACTACATTTAGGCTACGAGCGTACATGAACAC 4954
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGGCTCCCGTGTGCCAGGACCACTTGAAATTTGGGAGGGCGCTTTTACAGGCTC 5014
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5015 ACTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 5074
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 5075 CTGCTAGCGTACCAAGCCACCGTGTGCGCTAGGCGCTCAAGCCCTCCCGCATCGTGGAC 5134
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGGAAGTGTGATTTGGCTCAAGCCCACTTCAATGGGCCAACACCCCTGCTA 5194
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGACTGGGCGCTGTTTCAAGTGAATCACTCCCTGAGCAGCCAGTCAACCAATACATC 5254
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrThrThrTrpValLeuValGlyGly 640
Db 5255 ATGACATGCGATGTCCGCGCACCTGAGGTGCTCACGAGCACCTGGGTGCTGTTGGCGGC 5314
Qy 641 ValLeuAlaLeuAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTGTGCTGTTGGCGCGATTTGCTGTGCAACAGGCTCGGTGGTCATAGTGGCGAGG 5374
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTCTGTGTTTGGGAGCGCGGCAATCATACCTGACAGGGAGTCTCTTACCGAGAGTTC 5434
Qy 681 AspGluMetGluGluCys 686
|||||

Db 5435 GATGAGATGGAGAGTGC 5452
RESULT 4
US-07-910-760-9
; Sequence 9, Application US/07910760
; Patent No. 5683864
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07910,760
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9401 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 342..9374
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 372
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 867
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1341
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2148
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ile."
; FEATURE:
; NAME/KEY: misc_feature


```
/ LOCATION: 2883
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Asn."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3681
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3690
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4157
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Leu."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4323
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4701
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Tyr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4752
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 5970
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6183
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be His."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6186
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Cys."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6402
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7386
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7494
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7497
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ala."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7845
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/
```

```
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 8409
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9102
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9327
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Pro."
/ US-09-910-760-9
/
Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0
/
US-09-930-591-2 (1-686) x US-07-910-760-9 (1-9401)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3417 GTGGCGCCCATCACGGCGTACGCCACAGAGAGAGGGGCTCTTAGGGTGCAATACACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 3477 AGCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGCACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 3537 GCCCAACCTTCTGTGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCACGGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 3597 GGAACGAGGACCATCGCTGACCCAGGGTCTCTCATCCAGATGATACCAATGATAGAC 3656
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3657 CAAGACCTTGTGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCCTGCACTGC 3716
QY 101 GlySerSerAspLeuTyrIleuValThrArgHisAlaAspValIleProValArgArg 120
DB 3717 GGCTCTCGGACCTTTACCTGTGTACAGGACGCGCATGTCTCCGTCGCGCGCG 3776
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleuLysGlySerSer 140
DB 3777 GGTGATAGAGGGGCGACCTGCTGTCGCCCGGCCCATTTCTCTACTGTGAAGAGCTCCTCG 3836
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3837 GGGGGTCCGCTGTTGTGCCCCCGGGGCGACGCGTGGGGCATATTTAGGGCCGCGGTGTC 3896
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 3897 ACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGTGAGAACCTAGACACCAATG 3956
QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
DB 3957 AGTCCCCCGGTGTTACCGATAACTCTCTCCACAGTAGTGTGCCCCAGAGCTTCCAGGTG 4016
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 4017 GCTCACTCCATGCTCCACAGGACGCGGCAAAAGCAACCAAGGTCCCGGTGCATATGCA 4076
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
/
```

Db 4077 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGGCTTTGGT 4136
Qy 241 AlaTyrMetSerLysAlaHisGlyLeuAspProAsnIleArgThrGlyValArgThrIle 260
Db 4137 GCTTACATGTCCAAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 4196
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4197 ACCACTGGACGCCCATCAGCATCTCCACCTACCGCAAGTTCTTGGCCGACGGGGGTGC 4256
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4257 TCGGGGGGCGCTTATGACATAATAATTGTGACAGGTGCCACTCCAGGATGCCACATCC 4316
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4317 ATCTTGGGCATCGGCATGCTCTTGACCAAGCAGAGACTGGGGGGCGAGACTGTTGTG 4376
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4377 CTCGCCACCGCCACCCCTCGGGCTCCGTACTGTGCCCATCCCAACATCGAGAGGTT 4436
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4437 GCTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGCTATCCCTCCGAAAGTATC 4496
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4497 AAGGGGGGAGACATCTCATCTCTGTCATTCAAGAAGAAGTCGCGAGCACTCGCCGCA 4556
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4557 AAGCTGTGCGATTGGGCATCAATCGCGTGGCCCTACTACCGCGTCTTGACGTCCGTC 4616
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaThrMetThrGlyPheThr 420
Db 4617 ATCCGACACCGCGGATGTTGCTGTGGCAACCGATGCCCTCATGCCGCTATACC 4676
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4677 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTCGATTCAGC 4736
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4737 CTTGACCCCTACCTTCAACATTGAGACATACGCTCCCGAGGATGCTGCTCCCGCACT 4796
Qy 461 GluArgArgGlyArgThrGlyArgGlyPheProGlyIleTyrArgPheValAlaProGly 480
Db 4797 CAACGTGGGGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTGTGGCACCGGGG 4856
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4857 GAGCGCCCTCCGGCATGTTGCACTCGCTCGTCTCTGTGAGTGTATGACGAGGCTGT 4916
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4917 GCTTGGTATAGCTACGCGCCCGCAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 4977 CCGGGGCTTCCCGTGTCCAGGACCATCTTGAATTTTGGGAGGGCGCTTTTACAGGGCTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5037 ACTCATATAGTCCCACTTCTATCCAGAACAGAGAGTGGGGAGAACCTTCTTAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 5097 CTGGTAGCGTACCAAGCACCGTGTGCGTAGGCTCAAGCCCTCCCCCATCGTGGGAC 5156
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5157 CAGATGTGGAAGTGTGTTGATTTCGCTCAAGCCCAACCCCTCCATCGGCGCAACACCCCTGCTA 5216

Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5217 TACAGACTGGCGCGTGTTCAGAAATGAATCACCCTGACGACCCAGTCACCAATACATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5277 ATGACATGCATGTGCGCGCCACCTGGAGGTGTCACGAGCACCTGGGTGCTGTTGGCGGC 5336
Qy 641 ValLeuAlaAlaLeuAlaIleTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5337 GTCTGGCTGTTGGCGCGTATTGCTGTCAACAGCTGCGGTGTCATAGTGGCAGG 5396
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTGCTCTGTGCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGATGGAAGAGTGC 5474

RESULT 5
US-08-440-519-9
; Sequence 9, Application US/08440519
; Patent No. 5712087
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,519
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9401 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURES:
; NAME/KEY: CDS
; LOCATION: 342..9374
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 366
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Arg."

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 372
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 867
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1341
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ile."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402

```

```

OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also by Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-08-440-519-9

Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0

US-09-930-591-2 (1-686) x US-08-440-519-9 (1-9401)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3417 CTGGGCCCATCACGGCGTACGCCAGCAGACAGAGGGGCTCTTAGGGTGCATAATCACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCCTAACTGGCGGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3537 GCCCAACCTTCTGTGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCAACGGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGNACGAGGACCATCGCTCACCAAGGGTCTCTGTATCATCATGATATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTrpProLysGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGTGGGCTGGCCGCTCCGCAAGGTAGCGCTCATTCATTGACACCTGCACTGC 3716
QY 101 GlySerSerAspLeuTyrIleuValThrArgHisAlaAspValIleProValArgArg 120
Db 3717 GGCTCTCGGACCTTTTACCTGTGTCAGGAGCGCGCATGTCTATCCCGTGGCGCGCGG 3776

```

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlySerSer 140
DB 3777 GGTGATACAGAGGGGAGCTGCTGTCGCCCGGCCCATTTCTACTAGAAAGGCTCTCTCG 3836
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3837 GGGGGTCCGCTGTTGTGCCCCGGGGGACGCCGTGGGCATATTTAGGGCCCGGGTGGC 3896
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 3897 ACCCGTGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 3956
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3957 AGGTCCCCGGTGTTCACGGTAACCTCTCCACCCAGTAGTCCGCCAGAGCTTCACGGTG 4016
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 4017 GCTACCTTCCATGTCTCCACAGGACGCCGAAAGCACCAAGGTCCTCCGGCTGCATATGCA 4076
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 4077 GCTCAGGCTATAGGTCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 4136
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 4137 GCTTACATGTCCAAGGCTCATGGGATCGATCCTAACATCAGGACCGGGGTGAGAACAT 4196
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyLysCys 280
DB 4197 ACCACTGGCAGCCCATCAGCTACTCCACCTACCGCAAGTTCCTTGGCCGACGGGGTGC 4256
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 4257 TCGGGGGCGGCTTATGACATATATTTGTGACAGATGCCACTCCACGATGCCACATCC 4316
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 4317 ATCTTGGGATCGGCACCTGTCTTGACCAAGCAGAGACTCGGGGGGGAGACTGTTGTG 4376
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 4377 CTCCGCCACCGCCATCCCTCGGGGTCCGTCACTGTGCCCATCCCAACATCGAGAGGTT 4436
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 4437 GCTCTGTCCACCGGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAAGTATC 4496
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 4497 AAGGGGGGAGACATCTCATCTTGTCAATTCAAAGAAGAGTGGAGCACTCCCGCA 4556
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 4557 AAGCTGTGTGCTTGGGCATCAATCGCGCTACTACCGCGTCTTACGCTGCCGTC 4616
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 4617 ATCCCGACCGCGCATGTTGTGCTGGCAACCGATGCCCTCATGACGGGTATACC 4676
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 4677 GGCAGCTTCGACTGGGTATAGATGCAATACGTGTGTACCCAGACAGTCGATTTGAGC 4736
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 4737 CTTGACCTTACCTTACCATTTAGACAATCAGCTCCCGAGGATGCTGTCTCCCGACT 4796
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 4797 CAACGTGGGCGAGAGCTGGCAGGGGAAGCAGGCATCTACAGATTTGTGGCACCGGG 4856
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500

DB 4857 GAGCGCCCTCCGCGCATGTTGAGCTCGTCCGTCTCTGTGAGTGTATATGACGAGGCTGT 4916
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 4917 GCTTGGTATAGCTACGCCCGCCGAGACATACAGTTAGGCTACGAGCGTACATGAAACAC 4976
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
DB 4977 CCGGGGCTTCCGCTGTGCCAGGACCATCTTGAAATTTTGGGAGGGCGTCTTTACAGGCTC 5036
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 5037 ACTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 5096
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
DB 5097 CTGTAGCGTACCAAGCACCGGTGCTGCTAGGCTCAAGCCCTTCCCCCATCGGGAC 5156
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 5157 CAGATGTGAAGTGTGTTGATTCGCTCAAGCCACCTCCATGGGCCAACACCCCTGCTA 5216
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 5217 TACAGACTGGGCGCTGTTTCAATGAAATCACCCTGACGCCACCCAGTACCAAAATACATC 5276
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
DB 5277 ATGACATGCATGTCCGGCCGACCTGAGGTGCTGACGAGCACCTGGGTGCTCGTTGGCGGC 5336
QY 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyValArg 660
DB 5337 GTCTGTGCTGTTGGCGCGGTATTTGCTGCAACAGGCTGCGTGTGTCATAGTGGGCGG 5396
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 5397 GTCTGCTGTTCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 5456
QY 681 AspGluMetGluGluCys 686
DB 5457 GATGAGATGGAAGAGTGC 5474
RESULT 6
US-08-440-549-9
; Sequence 9, Application US/08440549
; Patent No. 6312889
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,549
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992

Db 3417 CTGGCGCCCATCGGCGGTACGCCAGCAGACAGAGGGCGCTCCTAGGCTGCATATACACC 3476
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCCTAACTGGCCCGGACAAAACCAAGTCGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3536
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
Db 3537 GCCAAACCTTCTTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACACACGGGGCC 3596
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGAACGAGGACCATCGCTCCACCCAGAGGTCTCTGTATCCAGATGTATACCAATGTAGAC 3656
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGTGGCTGGCGCGCTCCGCAAGGTAGCCGCTCATTTGACACCCCTGCACCTGC 3716
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 3717 GGCTCCTCGGACCTTTACCTGGTCAACGAGGACCGCGATGTCAATTCCTGGTGGCGCGG 3776
Qy 121 GlyAspGlyArgGlySerLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3777 GGTGATAGCAGGGGAGCGCTGCTGTGCGCCCGGCCCATTTCTACTTGAAGGCTCTCTCG 3836
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3837 GGGGGTCCGCTGTTGTGCGCCCGGGGCAACCGCTGGGCATATTTAGGGCGCGGTGTGC 3896
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3897 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCTCTGTGGAGAACCTAGAGACAAACATG 3956
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3957 AGTCCCGGTTGTCAAGGAGTCACTCAACCCCTCTGTCTCCACAGTAGTGGCCAGAGTCTCAGGTG 4016
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 4017 GCTCACCTCCATGCTCCACAGGCGAGCGGCAAAAGCACCAAGTCCCGGCTGCATATGCA 4076
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4077 GCTCAGGGCTATAAGGTAGTAGTACTCAACCCCTCTGTGTGTCACACACTGGGCTTTGGT 4136
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAsnProAsnIleArgThrGlyValArgThrIle 260
Db 4137 GCTTACATGTCCAAAGGCTCATGGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAAT 4196
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4197 ACCACTGGACGCCCATCATCGTACTCCACCTACCGGAAGTTCCTTGGCGAGCGGGGTGC 4256
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4257 TCGGGGGCGCTTATGACATAATAATTTGTGACAGTGCACCTCCACGGATGCCACATCC 4316
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4317 ATCTTGGGCACTCGGCATGCTCTTGACCAAGCAGAGACTCGGGGGCGGAGACTGGTTGTG 4376
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4377 CTGCGCACCGGCCACCCCTCCGGGCTCGTCACTGTGCCCCATCCCAACATCGAGAGAGTT 4436
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4437 GCTCTGTCCACCCAGGAGATCCCTTTTACGGCAAGGCTATCCCTCCCTCGAAGTAATC 4496
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4497 AAGGGGGGAGACATCTCATCTTCTGTCAATTCAAAGAGAAGTGCAGCAACTCGCGCA 4556

Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4557 AAGTCGTGCGATTGGGCATCAATGCCGTGCTTACTACCGCGTCTTTCAGTGTCCGTC 4616
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4617 ATCCCGACCAACGCGCATGTTGTCGTGTCGCAACCGATGCCCTCATGACCGGTATACC 4676
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4677 GCGGACTTCGACTCGGTGATAGACTGCATACGTGTGTACCCAGACAGTCGATTTGACG 4736
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4737 CTTGACCTTACCTTCCATTGAGACAATCAGGCTCCCCAGGATGCTGTCTCCGCACT 4796
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4797 CAACGTCCGGGCGAGGACTGGCAGGGGAAGCAGGCATCTACAGATTTGTGGCACCGGG 4856
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4857 GAGCGCCCTCCGGCATGTTGCACTCGTCCGTCTCTGTGAGTCTATGACGAGGCTGT 4916
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4917 GCTTGGTATGAGTCAACGCCCGCGAGACTACAGTAGGTAGGCTACGAGGTGATCAACACC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 4977 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGGAGGGCGCTTTTACAGGCTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5037 ACTCATATAGTCCCACTTTCATCCAGAAAGCAGAGTGGGAGAACCTTCTCTTAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 5097 CTGGTAGGTACCAAGCCACCGTGTGGCTTAGGCTCAAGCCCTCCCCCATCGTGGGAC 5156
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5157 CAGATGTGGAAGTGTGATTGCTCGCTCAAGCCACCTCCATGGGCCAACACCCCTGCTA 5216
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5217 TACAGACTGGGGCGCTGTTCAGAAATGAATCACCCTGACGCCACCCAGTCACCAATATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 5277 ATGACATGCATGTCCGCCGACCTGGAGGTGCTCAAGAGCACCTGGGTGCTCGTTGGCGGC 5336
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5337 GTCTGCGCTCTTGGCGCGTATTGCTGTCAACAGGCTCGTGGTGCATAGTGGGCGAG 5396
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTGCTCTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGATGGAAGAGTGC 5474

RESULT 7

US-08-823-895A-25

; Sequence 25, Application US/08823895A

; Patent No. 643159

; GENERAL INFORMATION:

; APPLICANT: Kevin P. Anderson

; TITLE OF INVENTION: Compositions And Methods For

; TITLE OF INVENTION: Treatment Of Hepatitis C Virus-Associated Diseases


```
Db 4857 GAGCCGCTCCGGGATGTTGACCTCGTCGTCCTCTGTGAGTGTATGACGAGCGCTGT 4916
Qy 501 AlaTpyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4917 GCTTGGTATGAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluValPheThrGlyLeu 540
Db 4977 CCGGGGCTTCGCGTGTGCAGACCATCTTGAATTTTGGGAGGGCGCTTTTACAGGCGTC 5036
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5037 ACTCATATAGATGCCACTTCTATCCAGAACAGAGAGTGGGGAGAACTTCCCTTAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db 5097 CTGTAGCTACCAAGCCACCGTGTGCGTAGGCGTCAAGCCCTCCCTCCCATCGTGGAC 5156
Qy 581 GlnMetTrpLysCysLeuLeuArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5157 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCACCTCCATGGGCCAACACCCCTGCTA 5216
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5217 TACAGACTGGGGCGCTGTTTCAGATGAATACCTCGACGACCCAGTCAACAAATACATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuGluValThrSerThrTrpValLeuValGlyGly 640
Db 5277 ATGACATGCATGCGCGGACCTGGAGTGTGTCAGAGACCTGGTGTCTGTTGGCGGC 5336
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValArg 660
Db 5337 GTCTGTGGTGTGTTGGCGCGCTATTGCTGTCAACAGGCTGCGTGTATAGTGGGAGG 5396
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTCTGTGTGTTGGGAGCGGCAATATCATCTGACAGGGAAGTCTCTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGTGGAGAGTGC 5474
```

RESULT 8

```
US-08-444-818-65
; Sequence 65, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/444, 818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403, 590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisha A.
```

```
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6785 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..6785
; US-08-444-818-65
Alignment Scores:
Pred. No.: 0 Length: 6785
Score: 3571.00 Matches: 671
Percent Similarity: 99.42% Conservative: 11
Best Local Similarity: 97.81% Mismatches: 4
Query Match: 98.70% Indels: 0
DB: 3 Gaps: 0
US-09-930-591-2 (1-686) x US-08-444-818-65 (1-6785)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1203 CTGGCGGCCATCATCGGGGTACGCCAGACAGAGGGGCTCTTAGGGTGCATATATCACC 1262
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Db 1263 AGCTAACTGCGCGGACAAAACCAAGTGAGGGTGGAGTCCAGATTGTGTCAACTGCT 1322
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 1323 GCCCAAACTTCTCTGGCAACGTCATCAATGGGGTGTGCTGGACTGTCTACACGGGGCC 1382
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1383 GGAACGAGAGACCATCGCGTCACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 1442
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1443 CAAGACCTTGTGGGCTGCGCGCTCCGCAAGGTAGCGCTCATTGACACCTCTGCACCTGC 1502
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 1503 GGCTCTCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCTATCCCGTGCAGCGGG 1562
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 1563 GGTGATAGCAGGGGCGACGCTGTCTCGCCCGGCCCATTTCTCTACTTGAAGGGTCTCTCG 1622
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 1623 GGGGGTCCGCTGTGTGTCGCCCGCGGGGACGCCGTGGGCATATTTAGGGCCGCGGTGTC 1682
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 1683 ACCCGTGGAGTGGCTAAGCGGTGAGCTTTATCTCTGTGGAGAACCTTAGACAAACCATG 1742
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 1743 AGGTCCCGGTGTTTCAGGATAACTCTCTCTCCACAGTAGTGGCCCCAGAGCTCCAGGTG 1802
Qy 201 AlahisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 1803 GCTCACCTCCATGCTCCACAGGAGCGGCAAGACCAAGCTCCCGGCTGCATATGCA 1862
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
```

Db 1863 GCTCAGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 1922
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 1923 GCTTACATGTCGAAGGCTCATGGATCGATCCTTAACATCAGGACCCGGGTGAGAACAT 1982
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 1983 ACCACTGGCAGCCCATCATGTAATCCACCTACGCAAGTTCCTTCCGACGCGGGTGC 2042
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2043 TCGGGGGCGCTTATGACATAATAATTTGTGACGACTGCCACTCCACGATGCCACATCC 2102
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2103 ATCTTGGGCATCGGCATGCTCTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTTGTG 2162
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2163 CTCGCCACCGCCACCTCCGGCTCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2222
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2223 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC 2282
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysValLysCysAspGluLeuAlaAla 380
Db 2283 AAGGGGGGAGACATCTCATCTTCTGTCATTCGAAGAAGAGTGCAGCAACTCGCGCA 2342
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2343 AAGCTGGTCGATGGGCATCATCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2402
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2403 ATCCGACCAACGCGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2462
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2463 GGGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTCGATTCAGC 2522
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 2523 CTGTGACCTTACCTTCCATTTGAGACATACGCTCCCCAGAGTGTCTCTCCGCACT 2582
QY 461 GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2583 CAACGTCGGGGCAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCAGGG 2642
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2643 GAGCGCCCTCCGGCATGTTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2702
QY 501 AlaTyrThrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2703 GCTTGGTATGAGTCAGCCCGCGGAGACTACAGTATAGGTACGAGGTCATGACACC 2762
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 2763 CCGGGGCTTCCGCTGCCAGGACCATCTTGAATTTTGGGAGGCGCTCTTTACAGGCGTC 2822
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 2823 ACTCATATGATGCCACTTTCATATCCAGAACAGAGTGGGGAGAACCTTCTCTTAC 2882
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 2883 CTGGTAGCTACCAAGCAGCGGTGGCTAGGGCTCAAGCCCTTCCCATCGTGGAC 2942
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2943 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCACCTCCATGGGCCCAACACCCCTGCTA 3002

QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3003 TACAGACTGGCGCTGTTTCAAGATGAATCAACCTGACGACCCAGTCCACCAATACATC 3062
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3063 ATGACATGATGTGCGGCCGACCTGGAGTCTGCACGAGCACCTGGGTGCTCGTTGGCGC 3122
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3123 GTCCTGGCTGCTTGTGGCCGCTATTCCTGTCAACAGGCTGCTGTGTCATAGTGGCAGG 3182
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3183 GTCGCTTCTGCGGGAAGCCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGATTTC 3242
QY 681 AspGluMetGluGluCys 686
Db 3243 GATGAGATGGAAGAGTGC 3260

RESULT 9

US-08-444-818-88
; Sequence 88, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8316 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8316
; US-08-444-818-88
Alignment Scores:
Pred. No.: 0 Length: 8316
Score: 3570.00 Matches: 671
Percent Similarity: 99.27% Conservative: 10
Best Local Similarity: 97.81% Mismatches: 5

US-08-444-818-122
 ; Sequence 122, Application US/08444818
 ; Patent No. 6150087
 ; GENERAL INFORMATION:
 ; APPLICANT: Chien, David Y.
 ; APPLICANT: Rutter, William J.
 ; TITLE OF INVENTION: NANV Diagnostics and Vaccines
 ; NUMBER OF SEQUENCES: 777
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Chiron Corporation
 ; STREET: 4560 Horton Street
 ; CITY: Emeryville
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94608-2916
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/444,818
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/403,590
 ; FILING DATE: 14-MAR-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Harbin, Alisa A.
 ; REGISTRATION NUMBER: 33,895
 ; REFERENCE/DOCKET NUMBER: 0110.002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (508)359-3876
 ; TELEFAX: (508)359-3885
 ; INFORMATION FOR SEQ ID NO: 122:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9185 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-444-818-122

Alignment Scores:
 Pred. No.: 0 Length: 9185
 Score: 3570.00 Matches: 671
 Percent Similarity: 99.27% Conservative: 10
 Best Local Similarity: 97.81% Mismatches: 5
 Query Match: 98.67% Indels: 0
 DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-122 (1-9185)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 3395 CTGCGGCCCATCATCGGCGTACGCCACAGCAAGAGGCGCTCTCCTAGGGTGCATAATCAC 3454
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 Db 3455 AGCCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCAGATTGTGTCAACTGCT 3514
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyVala 60
 Db 3515 GCCCAAACTTCTGGCAAGTGCATCAATGGGTGTGTGTAAGTCTTACCAAGGGGCC 3574
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 3575 GGAACGAGGACCATCGCGTCACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 3634
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 3635 CAAGACCTTGTGGCTGGCCCCGTCCGCAAGGTAGCGGCTCATGTAGACCCCTGCACTTGC 3694

QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 Db 3695 GGCTCTCTGGACCTTTTACCTGTGTACAGAGGACGCGCATGTCTATCCCTGCGCGCGG 3754
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleuLysGlySerSer 140
 Db 3755 GGTGATAGCAGGGGCGACCTGTCTGCGCCCGGCCCATTTCTTACTTTGAAAGCTCTCTCG 3814
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 3815 GGGGTCTCCCTGTGTGCCCCCGGGGACGCGGTGGGCATATTTAGGGCCGCGGTGTGC 3874
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
 Db 3875 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGTGAGAACCTCAGAGCAACCATG 3934
 QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
 Db 3935 AGGTCCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGTCCCGCAGAGCTTCCAGGTG 3994
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaValAla 220
 Db 3995 GCTCACTCCATGCTCCACAGGCGGCAAAAGCACCAAGGTCCCGCTGCATATGCA 4054
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 4055 GCTCAGGGCTATAAGGTGTCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTTGT 4114
 QY 241 AlaTyrMetSerIysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 4115 GCTTACATGTCCAAGGCTCATGGATCGATCTCTTAACATCAGGACCGGGGTGAGAACAT 4174
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 4175 ACACCTGGCAGCCCATACAGTACTCCACTACGCAAGTCTCTTGCCGACGCGGGTGC 4234
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 4235 TCGGGGGCGGTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCCACATCC 4294
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 4295 ATCTTGGGCATCGGCATCTCTTCAACAGCAGAGACTCGCGGGGCGAGACTGGTTGTG 4354
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 4355 CTCGCCACCGCCACCCCTCCGGGCTCCGTCATCTGTGCCCATCCCAACATCAGGAGGT 4414
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 4415 GCTCTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAAGTAATC 4474
 QY 361 LysGlyGlyArgHisIleIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
 Db 4475 AAGGGGGGAGACATCTCATCTCTGTCTATTCAAAGAAAGTGCAGAACCTCGCGCA 4534
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 4535 AAGCTGTCTGCATTTGGGCATCAATGCCGTGGCTACTACCGGGTCTTGACGTGTCCGTC 4594
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 4595 ATCCCGACCGCGCGATGTGTCTGTGTGGCAACCGATGCCCTCATGACCGGTATACC 4654
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 4655 GGCAGCTTCGACTCGGTGATAGACTGCATACGTGTGTCCACGACAGCTGATTTCAGC 4714
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 4715 CTGTACCTTACCTTACCATTGAGCAATCAGCTTCCCCCAGGATGCTGTCTCCGCACT 4774
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480

Db	4775	CAACGTCGGGGCAGGACTGGCAGGGGAAGCAGGCATCTACAGATTGTGGCCACCGGGG	4834
Qy	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	4835	GAGCGCCCTCCGGCATGTTCCGACTCGTCCGTCCTCTGTGNGTGCTATGACGACGGCTGT	4894
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	4895	GCTTTGGTATGAGCTCAGCGCCGCGAGACTACAGTTAGGCTACGAGCGGTACATGAACACC	4954
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
Db	4955	CCGGGGCTTCCCGTGTCCAGGACCATCTTGAAATTTTGGGAGGGCGTCTTTACAGGGCCTC	5014
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	5015	ACTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCCTTAC	5074
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp	580
Db	5075	CTGTAGCTAGTACCAAGCACCGTGTGGCTAGGGCTCAAGGCCCTCCCCCATCGTGGGAC	5134
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	5135	CAGATGTGAAGTGTGTGATTCGGCTCAAGCCCACTTCATGGGCCCAACACCCCTCGCTA	5194
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	5195	TACAGCTGGCGCTGTTCAGATGAATCACCTTCAGCCAGCACCCAGTCAACAAATACATC	5254
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	5255	ATCACATGCATGTCCGCGCAGCTGGAGTGGTCAACGAGCACCTGGGTGCTCGTTGGCGGC	5314
Qy	641	ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	5315	GTCTGTGCTCTTGGCCCGGTATFGCTGTCAACAGGCTGCGTGGTCAATGTTGGGACGG	5374
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	5375	GTGCTTGTTCGGGAAGCCGGCATCATACCTGACAGGGAGTCTCTTACCAGAGTTTC	5434
Qy	681	AspGluMetGluGluCys	686
Db	5435	GATCAGATGGAAGTGC	5452

RESULT 11

US-08-444--818-123/c	
/ Sequence 123, Application US/08444818	
/ Patent No. 6150087	
/ GENERAL INFORMATION:	
/ APPLICANT: Chien, David Y.	
/ APPLICANT: Ruter, William J.	
/ TITLE OF INVENTION: NANBV Diagnostics and Vaccines	
/ NUMBER OF SEQUENCES: 777	
/ CORRESPONDENCE ADDRESS:	
/ ADDRESSEE: Chiron Corporation	
/ STREET: 4560 Horton Street	
/ CITY: Emeryville	
/ STATE: CA	
/ COUNTRY: USA	
/ ZIP: 94608-2916	
/ COMPUTER READABLE FORM:	
/ MEDIUM TYPE: Floppy disk	
/ COMPUTER: IBM PC compatible	
/ OPERATING SYSTEM: PC-DOS/MS-DOS	
/ SOFTWARE: PatentIn Release #1.0, Version #1.30	
/ CURRENT APPLICATION DATA:	
/ APPLICATION NUMBER: US/08/444,818	
/ FILING DATE:	
/ CLASSIFICATION: 424	
/ PRIOR APPLICATION DATA:	

3351 CAGACCCCTTGGGCGCCTTCGCCAAGGTAGCCGCCTCATTTATTAACAACCCCTGCATTCG 3359
 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 5491 GGCCTCCTCGGACCCTTTACCTGGTCACGAGGCACCGCATGCTCATTCCTCGTGCCGCGG 5432
 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 5431 GGTCATAGCAGGGCAGCTGCTGTCGCCCGCGCCATTTCCTACTTGAAGAGCTCTCTCG 5372
 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 5371 GGGGGTCGCTGTTGTCCCCCGGGGCACGCGTGGGCATATTATAGGCGCGCGGTGTGC 5312
 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 5311 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATTCCTGTGGAGAACCTTAGNAGACCAATG 5252
 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSertyrGlnVal 200
 5251 AGGTCCCCCGGTGTTTCACGGATAACTCTCTCCACCACTAGTGTGCCCCCAGAGCTTCCAGGTG 5192
 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 5191 GCTCACCTCCATGTCTCCACAGGCAGCGGCAAAAGCACCAGGTCCCGGTGCATATGCA 5132

```
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 5131 GCTCAGGCTATAAGGTGTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGCT 5072
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 5071 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAACAATT 5012
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyPheLeuAlaAspGlyGlyCys 280
Db 5011 ACCACTGGCAGCCCATCAGTACTCCACTACGGCAAGTTCCTTCCGACGCGGGGTGC 4952
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4951 TCGGGGGCGGTTATGACATAAATTTGTGACGAGTGCACATCCACGGATGCCACATCC 4892
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4891 ATCTTGGGCAATCGGCATCTGCTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTTGTG 4832
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4831 CTGCGCACCGCCACCCCTCGGGGCTCGGTCACTGTGCCCCATCCCAACATCGAGGAGGT 4772
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4771 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTAATC 4712
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4711 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAAGAAGTGCAGCAACTCGCGCA 4652
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrArgGlyLeuAspValSerVal 400
Db 4651 AAGTGTGTCGATTTGGGCATCAATCGCGTGGCTACTACCGCGGTCTTGACGTCCGTC 4592
QY 401 IleProThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4591 ATCCGACACCGCGGATGTTGTCGTGCGCAACCGATGCCCTCATGACCGGCTATACC 4532
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4531 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCAACCCAGACAGTCGATTTTCAGC 4472
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4471 CTGACCCCTACCTTCACATTTAGACAAATCAGCTCCCCAGAGTGTCTCCGCACT 4412
QY 461 GlnArgArgGlyArgThrGlyArgGlyPheProGlyIleTyrArgPheValAlaProGly 480
Db 4411 CAACGTCGGGCGAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGG 4352
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4351 GAGCGCCCTCCGCGATGTTTCGACTCGCTCGCTCTCTGTGAGTGTATGACGAGGCTGT 4292
QY 501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4291 GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTACGAGGCTACATGAACACC 4232
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrPheGluGlyValPheThrGlyLeu 540
Db 4231 CCGGGGCTTCCCGTGTCCAGGACCATCTTGAATTTTGGAGGGCGTCTTTACAGGCGTC 4172
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 4171 ACTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGGAACCTTCCTTAC 4112
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580
Db 4111 CTGGTAGGTACCAAGCAGCGTGTGCTAGGGCTCAGGCCCTCCCCCATCGTGGGAC 4052
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
```

```
Db 4051 CAGATGTGGAAGTGTGTTGATTGCGCTCAAGCCCAACCCCTCCATGGGCCAACACCCCTGCTA 3992
QY 601 TyrArgLeuGlyAlaValAlaGlnAsnGluValThrLeuThrHisProValThrLysThrIle 620
Db 3991 TACAGACTGGGGCGCTGTTTCAAGATGAATCAACCTTGACGACCCAGTACCAGATACATC 3932
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3931 ATGACATGCATGTCCGCGACCTGGAGGTGCTGACAGCACCTGGGTGCTGTTGGCGGC 3872
QY 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValIleValGlyArg 660
Db 3871 GTCTGTGGTGTCTTGGCGCGTAITTCCTGTCAACAGGCTGCGTGTATAGTGGCGAG 3812
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3811 GTGCTGTTGTCCGGGAAGCGGCAATCATACCTGACAGGAAGTCTCTTACCAGAGTTC 3752
QY 681 AspGluMetGluGluCys 686
Db 3751 GATGAGATGAAGAGTGC 3734
RESULT 12
US-09-881-239-2
; Sequence 2, Application US/09881239
; Patent No. 6630298
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARANGEL, Phillip
; APPLICANT: TANDESKE, Laura
; APPLICANT: GEORGE-NASCIMENTO, Carlos
; APPLICANT: COIT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY
; FILE REFERENCE: 2302-16073 / PP16073.003
; CURRENT APPLICATION NUMBER: US/09/881,239
; CURRENT FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: representative NS3/4a conformational antigen
; NAME/KEY: CDS
; LOCATION: (1)..(2058)
US-09-881-239-2
Alignment Scores:
Pred. No.: 0 Length: 2058
Score: 3565.00 Matches: 671
Percent Similarity: 99.13% Conservatives: 9
Best Local Similarity: 97.81% Mismatches: 6
Query Match: 98.54% Indels: 0
DB: 4 Gaps: 0
US-09-930-591-2 (1-686) x US-09-881-239-2 (1-2058)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1 ATGGCGCCCATCACGGCGTACGCCAGCAGACAAAGGGGCTCTCCTAGGGTGCATAATCACC 60
QY 21 SerLeuThrGlyArgAspLysGlnValGluGlyValGlnIleValSerThrAla 40
Db 61 AGCTTAACATGGCGGGGACAAAACCAAGTGGAGGGTGAAGTCCAGATTGTGTCACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 121 GCCCAACCTTCTCTGCAACGTCATCANTGGGTGTGTGGACTGTCTTACCACGGGGCC 180
```

Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 181 GGAACAGGACCATCGCGCTCAACCAAGGGTCTGTGCATCCAGATGTATACCAATGTAGAC 240
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 241 CAAGACCTTGGCGTGGCGCTCCGCAAGGTAGCCGATCATTTGACACCTCGCATTCG 300
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 301 GGCTCTCGGACCTTACCTGGTCAGAGGACCGCGATGTCATTCCTGGCGCGGG 360
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 361 GGTGATAGCAGGGGAGCGCTGCTGTGCGCGCGGCCATTTCTACTTGAAGGCTCTCTCG 420
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 421 GGGGTCGCGCTGTGTGCGCGGGGACCGCGTGGGCATATTTAGGGCGCGGTGTGC 480
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 481 ACCGCTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACCAACATG 540
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 541 AGGTCCCGGTGTTCACGGATAACTCTCTCCACCCAGTAGTGGCCCGCAGAGCTCCAGGTG 600
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 601 GCTCACTTCCATGTCCACAGGAGCGGCAAAAGCACCAAGGTCCCGGTGTCATATGCA 660
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 661 GCTCAGGGCTATAGGTCTAGTACTCAACCCCTCTGTGTGTGCAACTGGGCTTTGGT 720
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 721 GCTTACATGTCCAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAT 780
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyLysCys 280
Db 781 ACCACTGGCAGCCCATCATCATCTCCACCTACCGCAAGTTCTTGGCGAGCGGGGTGC 840
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 841 TCGGGGGCGGCTTATGACATATATTTGTGACAGGTGCCACTCCAGATGCCACATCC 900
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 901 ATCTTGGCATTTGGCACTGTCTTGACCAAGCAGAGACTGGCGGGCGGAGACTGGTTGTG 960
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 961 CTGCGCCACCGCACCCCTCCGGGTCTCGTCACTGTGCCCATCCCAACATCGAGGAGTT 1020
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1021 GCTCTGCCACCGGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAACT 1080
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 1081 AAGGGGGGAGACATCTCATCTCTGTGTCATTCAAAGAAGAGTGGAGCAACTCGCCGCA 1140
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 1141 AAGTGTGTGTCATTTGGGCATCAATCCGCTGGCTACTACCGCGTCTTTCAGTGTCCGTC 1200
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 1201 ATCCCGGCCATTCGCGCATTTGTGTGTGTCGCAACCGATGCCCTCATGACCGGTATACC 1260
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

Db 1261 GGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTGCATTTTCAGC 1320
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTTCAGCCCTACCTTCACCATTTGACAAATCAGCTCCCCCAAGATGCTGTCTCCGCACT 1380
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 1381 CAACGTCCGGGCGAGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGG 1440
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 1441 GAGCGCCCTCCCGCATGTTTCGACTCGTCCGCTCTCTGTAGTGTATGACCGAGGTGT 1500
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 1501 GCTTGGTATGAGTTCACGCCCGCGAGACTACAGTTAGGCTTACGAGCGTATACGAACAC 1560
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 1561 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGGAGGGCGTCTTTACAGGCTC 1620
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 1621 ACTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 1680
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 1681 CTGGTAGCGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCTCCCCCATCGTGGAC 1740
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGGAGTGTGTGATTCGCTCAAGCCACCTCTCATATGGGCCAACACCTCTGCTA 1800
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TACAGACTGGCGCTGTTCAGATGAATCACCCTGACGACCCAGTCCACCAATACATC 1860
Qy 621 MetThrCysMetSerAlaAspLeuValValThrSerThrTyrValLeuValGlyGly 640
Db 1861 ATGACATGCATGTGCGCGGACCTGGAGGTCTCAGGAGCACCTGGGTGCTGTGGCGGC 1920
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTCTGTGCTGTTGGCGGCTATTCCTGTCAACAGGCTGCGGTGCTATAGTGGGAGG 1980
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 GTGCTCTTGTCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 2040
Qy 681 AspGluMetGluGluCys 686
Db 2041 GATGAGATGGAAGTGC 2058

RESULT 13

US-09-981-654-1
; Sequence 1, Application US/09881654
; Patent No. 6632601
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANGEL, Phillip
; APPLICANT: TANDESKE, Laura
; APPLICANT: GEORGE-NASCIEMENTO, Carlos
; APPLICANT: COLT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / P17039.002
; CURRENT APPLICATION NUMBER: US/09/881,654
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811


```
/ PRIOR FILING DATE: 2001-04-02
/ PRIOR APPLICATION NUMBER: 60/280,867
/ PRIOR FILING DATE: 2001-04-02
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: Patent in Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 2058
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:
/ OTHER INFORMATION: representative NS3/4a conformational antigen
/ NAME/KEY: CDS
/ LOCATION: (1)..(2058)
US-09-881-654-1

Alignment Scores:
Pred. No.: 0 Length: 2058
Score: 3565.00 Matches: 671
Percent Similarity: 99.13% Conservatives: 9
Best Local Similarity: 97.81% Mismatches: 6
Query Match: 98.54% Indels: 0
DB: 4 Gaps: 0

US-09-930-591-2 (1-686) x US-09-881-654-1 (1-2058)

QY 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 1 ATGGCGCCCATCAGCGGTACGCCACAGACAGAAAGGGGCTCTTAGGGTGCAATACAC 60
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 61 AGCTTAATGCGCGGACAAACCAAGTGGAGGGTGAGTCCAGATTGTGCACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 121 GCCCAAACTTCTCGCAACGTCATCAATGGGGTGTGCTGGACTGTGTACACCGGGCC 180
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 181 GGAACGAGGACCATCGCGTCACCAAGGGTCCCTGTCTATCCAGATGTATACCAATGTAGAC 240
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100
DB 241 CAAGACCTTGTGGCTGGCGCTCCGCAAGTAGCCGATCATGACACCTCGACTGC 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 301 GGCTCTCGGACCTTTACCTGGTCACGAGGACGCGCATGTCATTCCTCGCGCGCGG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 361 GGTGATAGAGGACGCGCTGTGTCGCCCGGCCCATTTCTTACTTTGAAGGCTCTCTCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 421 GGGGTCCGTGTGTGTCGCCCGGGGACGCCGTGGGCATATTTAGGGCGCGGTGTC 480
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCTCTGTGGAGAACCTTAGAGACAACCATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 541 AGGTCCCGGTGTTCAGGATAACTCTCTCCACAGTAGTGCCCGGAGAGCTTCCAGGTG 600
QY 201 AlahisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 601 GCTCACCTCCATGCTCCACAGGACGCGGCAAAAGCACCAAGTCCCGGTGTCATATGCA 660
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 661 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTGTCGAACACTGGGGCTTGGT 720
```

```
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTCAGAGCTCATGGATCGATCTTAACATCAGACCGGGGTGAGAACAAAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 781 ACCACTGGCAGGCCCATCATCAGTACTCACCTACGCAAGTTCCTTGGCCACGCGCGGTGC 840
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 841 TCGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGATGCCATCC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 901 ATCTTGGGCATTTGGCACTGTCTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 961 CTCGCCACGCCACCCCTCGGGCTCCGTCACTGTGCCCCCATCCCAACATCGAGGAGGT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 1021 GCTCTGTCCACCACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC 1080
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 1081 AAGGGGGGAGACATCTCATCTCTGTCTTCAAGAAGAGTGCAGCAACTCGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 1141 AAGCTGTCGCATTTGGGCATCAATGCGTGGCTACTACCGCGGTCTTGACGTGTCGTC 1200
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCCGCCCATCGGCATGTTGTGTCGGCAACCGATGCCCTCAAGCCGCTATACG 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 1261 GCGCACTTCGACTCGGTGATAGACTGCAATAGTGTGTCAACCCAGACAGTCGATTTT 1320
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCTTACCTTCACTTCCATTTGAGACAATCAGCTCCCGCCAGATCTGTCTCCGC 1380
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1381 CAACGTGGGGGACGAGTGGCAGGGGAGCCAGGCACTTACAGATTTGTGGCACCGGGG 1440
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 1441 GAGCGCCCTCCGGCATGTTGCACTCGTCTCTGTGAGTGTATGACGCGAGGTGT 1500
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 1501 GCTTGTATGAGTCAACGCCCGCGAGACTACAGTTAGCTACGAGCTACATGAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
DB 1561 CCGGGGCTTCCCGGTGCGCAGGACCATCTTGAATTTTGGAGGGCGCTTTTACAGGCTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 1621 ACTCATATAGTCCCATTTCTATCCAGACAAGCAGAGTGGGAGAACCTTCCTTAC 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
DB 1681 CTGGTAGCGTACCAAGCCACCGGTGCTGCTAGGGCTCAAGGCCCTCCCCCATCGTGGAC 1740
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 1741 CAGATGTGGAAGTGTGTGATTCGCTTCAAGCCCAACCTCTCATGGGCCAACACCCCTGCTA 1800
```

QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIleTyr 620
Db 1801 TACAGACGGCGCTGTTCAGAAATGAAATACCTCGACGACCCAGTCACCAATATACATC 1860
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATGACATGCTATGTCGGCGGACCTGGAGTGTCTGACGACACCTGGTGTCTGTTGGCGGC 1920
QY 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTCTGGCTGCTTGGCGCGCTATTGCTGCTCAACAGCGCTGGTGTCTAGTGGCGAGG 1980
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 GTCTGTTGTCGGGAAGCGGCAATCATCTGACAGGAAGTCTCTACCGAGAGTTC 2040
QY 681 AspGluMetGluGluCys 686
Db 2041 CATGATGGAGAGTGC 2058

RESULT 14

US-10-637-323-1

; Sequence 1, Application US/10637323

; Patent No. 6797809

; GENERAL INFORMATION:

; APPLICANT: CHIEN, David Y.

; APPLICANT: ARCANGEL, Phillip

; APPLICANT: TANDESKE, Laura

; APPLICANT: GEORGE-NASCIEMENTO, Carlos

; APPLICANT: COLT, Doris

; APPLICANT: MEDINA-SELBY, Angelica

; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES

; FILE REFERENCE: 2302-17039 / PP17039.002

; CURRENT APPLICATION NUMBER: US/10/637,323

; CURRENT FILING DATE: 2003-08-08

; PRIOR APPLICATION NUMBER: US/09/881,654

; PRIOR FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: 60/212,082

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: 60/280,811

; PRIOR FILING DATE: 2001-04-02

; PRIOR APPLICATION NUMBER: 60/280,867

; PRIOR FILING DATE: 2001-04-02

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 2058

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: representative NS3/4a conformational antigen

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2058)

US-10-637-323-1

Alignment Scores:

Pred. No.:	0	Length:	2058
Score:	3565.00	Matches:	671
Percent Similarity:	99.13%	Conservative:	9
Best Local Similarity:	97.81%	Mismatches:	6
Query Match:	98.54%	Indels:	0
DB:	4	Gaps:	0

US-09-930-591-2 (1-686) x US-10-637-323-1 (1-2058)

QY 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuGlyCysIleIleThr 20
Db 1 ATGGCGCCCAATCAGCGGTACGCCAGACAAAGGGGCTCTAGGGTGCATTAATCACC 60
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerAla 40

Db 61 AGCCTAACTGGCCCGGCAAAACCAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 121 GCCCAAACTTCTGGCAACGTCATCAATGGGGTGTCTGGACTGTCTACACCGGGGCC 180
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 181 GGAACGAGGACCATCGGCTCACCCAGAGGTCCTGTCTCATCCAGATGTATACCAATGTAGAC 240
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 241 CAAGACCTTGGGCTGGCCGCTCCGCAAGTAGTACCGATCAATTGACACCTCGACATTGC 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 301 GGCTCCTCGGACCTTACCTGGTCACGAGGACGCGCATTCATTCCTGCGCGGGCGG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 361 GGTGATAGAGGGGACGCTGTCTGCGCCCGGCCCATTTCTACTTTGAAAGGCTCCTCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 421 GGGGTCCTGCTGTTCGCCCGCGGGCACGCCGTGGGCATATTTAGGGCGCGGTGTGC 480
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 481 ACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCTGTGGAGAACCTTAGAGAACCAATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 541 AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTCCCGAGAGGTTCCAGGTG 600
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 601 GCTCACCTCCATGCTCCACAGGCGAGGCAAAAGCACCAAGGTCCCGGTGCATATGCA 660
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 661 GCTCAGGGCTATAGGTGTCTAGTACTCAACCCCTCTGTGTGCAACACTGGGCTTTGGT 720
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 721 GCTTACATGTCCAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 781 ACCACTGGCAGCCCATCATCGTACTCCACCTAGCGCAAGTTCTCTGCGACGCGGGTGC 840
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 841 TCGGGGGCGCTTATGACATAATAATTGTGAGAGTGCCACTCCACGGATGCCACATCC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 901 ATCTTGGCAATGGCACTGCTTGTACCAAGACAGAGACTCGGGGGGAGACTGGTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 961 CTGCGCACCGCCACCCCTCGGGCTCGTCACTGTGCCCCCATCCCAACATCGAGAGGTT 1020
QY 341 AlaIleSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1021 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCCTCGAAGTATC 1080
QY 361 LysGlyGlyVArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 1081 AAGGGGGGAGAGATCTCATCTCTTGTCAATCAAAAGAAAGTGCAGCAACTCGCCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 1141 AAGCTGCTCGATTGGGCAATCAATCGCTGCGCTACTACCGGCTCTTGACGTGCTCGTC 1200

```
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 1201 ATCCGCGCCATCGCGATGTTGCTGCTGCGCAACCGATGCCCTCATGACCGGCTATACG 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 1261 GCGGACTTCGACTCGGTGATAGACTGCAATACATGCTGTACCCAGACAGCTGATTTTCAGC 1320
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTTGACCTTACCTTACCATTTAGACATATCACCTCTCCCAAGATGCTGTCTCCGCACT 1380
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 1381 CAACGTCGGGCGAGGACTGCGCGGGAAGCCAGGCACTACAGATTTGTGGCAACCGGG 1440
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 1441 GAGCGCCCTTCGCGCATGTTTCGACTCGTCCGCTCTCTGTGAGTGTATGACGCGAGGCTGT 1500
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 1501 GCTTGTGTATGAGTTCACGCGCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 1561 CCGGGGCTTCCTGTCGCGAGGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCGCTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 1621 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 1681 CTGGTACGTAACCAAGCCCGTGTGCGCTAGGCGCTCAAGCCCTCTCCCATCGTGGGAC 1740
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGGAGTGTGTTGATGCTTCAAGCCCAACCTCTCATGGCCCAACACCCCTGCTA 1800
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TACAGACTGGCGCGTGTTCAGATGAATCAACCTCAGCCAGCCAGTCAACAAATACATC 1860
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATGACATGCAATGTCGCGCGACCTGGAGGTGCTCAGAGCACTGGGTGCTGTGTGGCGGC 1920
QY 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTCTGGCTGCTTGGCGCGTATTCCTCTCAACAGGCTGCTGCTCATAGTGGCGAGG 1980
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 GTGCTGTGTTCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 2040
QY 681 AspGluMetGluGluCys 686
Db 2041 GATGAGATGAAGAGTGTC 2058

RESULT 15
US-08-444-818-137
; Sequence 137, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
```

```
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 137:
SEQUENCE CHARACTERISTICS:
LENGTH: 8987 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..8985
US-08-444-818-137

Alignment Scores:
Pred. No.: 0 Length: 8987
Score: 3565.00 Matches: 671
Percent Similarity: 99.27% Conservative: 10
Best Local Similarity: 97.81% Mismatches: 5
Query Match: 98.54% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-137 (1-8987)
QY 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3076 CTGGCGCCCATCATCGCGGTACGCCAGCAGACAGAGGGGCTCTTAGGGTGCATAATCACC 3135
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3136 AGCTTAACCTGGCGGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 3195
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3196 GCCCAACCTTCTCGCAACGTGCATCAATGGGGTGTCTGGACTGTCTACCAACGGGGCC 3255
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3256 GGAACGAGGACCATCGCGTCACCAAGGCTCTGTCTCATCCAGATGTATACCATGTAGAC 3315
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3316 CAAGACCTTGTGGGCTGGCGCTCGCAAGGTAGCGGCTCATTCAGATGTATACCATGTGC 3375
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 3376 GGCTCTCTCGGACCTTTACTCTGGTACGAGGACACCGCGATGTCTATCCCGTCCCGCGCG 3435
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3436 GGATGATGACGAGGCGAGCTGTGTGCGCCCGCCCATTTCTTACTTGAANGCTCCTCG 3495
```

```
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3496 GGGGGTCCGCTGGTGTGTCGGCGGGGACCCGCGGGCATATTTAGGGCGCGGTGTGC 3555
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3556 ACCCGTGGAGTGGCTAAGCGCGGTGACCTTATCTCTGTGGAGAACCTTAGAGACAACCATG 3615
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3616 AGGTCCCGGTGTTCCAGGATAACTCTCTCCACCAGTAGTGCCCGCAGAGCTTCCAGGTG 3675
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 3676 GCTCACTCCATGCTCCACAGGAGCGGCAAGCAACAAAGGTCCCGGCTGCATATGCA 3735
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 3736 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 3795
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 3796 GCTTACATGTCACAGGCTCATGGATCGATCCTTAACATCAGACCGGGTGAGAACAAAT 3855
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 3856 ACCACTGGCAGCCCATCACGTACTCCACCTACGCGCAAGTTCTTTGGCGACGGGGGTGC 3915
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 3916 TCGGGGGCGCTTATGACATAATTAATTTGACGAGTGCCACTCCACGGATGCCACATCC 3975
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 3976 ATCTTGGGATCGGACGTGCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTG 4035
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4036 CTCCGCCACCGCACCCCTCCGGGTCGGTCACTGTGCCCCATCCCAACATCGAGAGGTT 4095
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4096 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 4155
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4156 AAGGGGGGAGACATCTCTCTGTCAATCAAGAAGAGTGGCAGCACTCGCCGCA 4215
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4216 AAGCTGGTTCGATTTGGGCATCAATGCGGTGGCTTACTACCGCGGTCTTGACGTCCGTC 4275
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4276 ATCCCGACCAAGCGGAGTGTGTCGTGGCAACCGATGCCCTCATGACCGGCTATACC 4335
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4336 GGCACCTTCGACTCGGTAGACATGCATACGTGTGTACCCAGACAGTCGATTTTCCAGC 4395
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4396 CTTGACCCCTACCTTCAACATTGAGACAATACGCTCCCGAGGATGCTGTCTCCGCACT 4455
Qy 461 GlnArgArgGlyArgThrGlyArgGlyPheProGlyIleTyrArgPheValAlaProGly 480
Db 4456 CAACGTGGGGCAGGACTGGCAGGGGAAGCAGGCATCAACAGATTGTGGCACCGGGG 4515
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4516 GAGCGCCCTCCGCGATGTTGACCTCGTCCGTCTCTGTGAGTGTATGACGAGGCTGT 4575
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
```

```
Db 4576 GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4635
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 4636 CCGGGGCTTCCGCTGTCCAGGACCACTTGAATTTTGGAGGGGCTTTTACAGGCCTC 4695
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 4696 ACTCATATAGATGCCACTTCTATCCAGACAAGCAGAGTGGGGAGAACCTTCTCTTAC 4755
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db 4756 CTGCTAGCGTACCAAGCACCGTGTGCGCTCAGGCTCAAGCCCTCCCCCATCGTGGGAC 4815
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 4816 CAGATGTGGAAAGTGTGTTGATTCGCTCAGCCCACTTCCATGGGCCCAACCCCTGCTA 4875
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 4876 TACAGACTGGGCGCTGTTTCAAGATGAATCACCCTGACGCAACCCAGTCACCAATATAC 4935
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 4936 ATGACATGTCATGTCCGGCCGACCTGGAGTGTCTACAGACACCTGGGTGCTGTTGGCG 4995
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 4996 GTCTGCTGCTTTTGGCGCGCTATTGCTGTCAACAGGCTGCGTGGTCCATAGTGGGAG 5055
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5056 GTGCTCTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAGTCTCTTACCGAGAGTTC 5115
Qy 681 AspGluMetGluGluCys 686
Db 5116 GATGAGATGGAAGAGTGC 5133
```

Search completed: February 27, 2005, 07:25:22
Job time : 412 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2005, 03:26:17 ; Search time 4728 Seconds
(without alignments)
5522.861 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPTAYAQTRGLLGLIIT.....PALIPDREVLVREDFEMEEC 686

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US930591/runat_25022005_143613_25686/app_query.fasta_1.839
-DB=EST -QFMT=fastap -SUFFIX=src -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US930591 @CGN 1.1 5180 @runat_25022005_143613_25686 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gse1.*
9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	181	5.0	CR591941 full-leng
2	163	4.5	BC046781 Mus muscu
3	160	4.4	CL974652 OsIFCC026
4	150.5	4.2	AK004665 Mus muscu
5	141	3.9	AY411078 Homo sapi
6	140.5	3.9	BU054791 UI-M-FD0-
7	136.5	3.8	BC059369 Homo sapi
8	133.5	3.7	CD359697 AGENCOURT
9	132.5	3.7	BU187274 AGENCOURT

10	131	3.6	949	1	AL560974	AL560974
11	129.5	3.6	2856	3	CR613419	full-leng
12	129	3.6	2972	3	AK031534	Mus muscu
13	129	3.6	2984	3	AK031679	Mus muscu
14	129	3.6	2986	3	AK028274	Mus muscu
15	129	3.6	3956	3	BC043699	Mus muscu
16	128.5	3.6	1697	9	AY404177	Mus muscu
17	128	3.5	636	7	CN788030	4122309 B
18	128	3.5	845	7	CV106671	AGENCOURT
19	127.5	3.5	691	5	BU054966	UI-M-FD0-
20	127.5	3.5	738	6	CD240900	AGENCOURT
21	127.5	3.5	2976	9	AY400284	Homo sapi
22	126	3.5	1283	5	BQ709745	AGENCOURT
23	126	3.5	1642	3	CR610486	full-leng
24	125.5	3.5	2484	3	CNS0AAU3	Arabidops
25	124.5	3.4	623	7	CF131593	UI-HF-F00
26	124.5	3.4	2388	9	AY418898	Mus muscu
27	124.5	3.4	3633	3	AY383690	Rattus no
28	124	3.4	1185	9	CU962117	OsIFCC007
29	124	3.4	2388	9	AY418896	Homo sapi
30	123.5	3.4	1788	3	AY105041	Zea mays
31	123.5	3.4	1818	9	CL973139	OsIFCC042
32	123	3.4	790	7	CK017540	AGENCOURT
33	123	3.4	3461	9	AY398774	Mus muscu
34	123	3.4	3802	3	AK004733	Mus muscu
35	123	3.4	4327	3	BC058331	Mus muscu
36	123	3.4	4640	3	BC062885	Mus muscu
37	122	3.4	3879	3	BC028405	Homo sapi
38	121.5	3.4	3211	3	BC041392	Homo sapi
39	121	3.3	2270	9	AY418897	Pan trogl
40	121	3.3	7320	3	CR627021	Homo sapi
41	120.5	3.3	478	2	BF725559	bx18a02.y
42	120.5	3.3	694	4	BI088407	602852901
43	120.5	3.3	3856	3	AK078552	Mus muscu
44	120	3.3	651	1	AJ729569	AJ729569
45	120	3.3	1143	5	BM926541	AGENCOURT

ALIGNMENTS

CR591941 2236 bp mRNA linear HTC 21-JUL-2004

full-length cDNA clone CSODL006YA04 of B cells (Ramos cell line)

Cot 25-normalized of Homo sapiens (human).

CR591941 GI:50472748

HTC; CNSLT_CDNA.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 2236)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished

Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue

2 (bases 1 to 2236)

Genoscope.

Direct Submission

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen.

Location/Qualifiers

1. .2236

/organism="Homo sapiens"

COMMENT

FEATURES

source


```
QY 396 euAspValSerValIleProThrSerGlyAspValValValValAlaThrAspAlaLeuM 416
Db 1895 TGCAGATCCAGCTGACCCAGCCCTCGGAGATATCTGGTG-----TTCC 1939
QY 416 etThrGlyPheThrGlyAspPheAspSerValIleAspCys-----429
Db 1940 TGACGGGACA-GGAGGAGATTGAGGCTGCTGTGAGATGCTCCAGGACCGCTCCGCGAG 1998
QY 429 -----429
Db 1999 CTGGGCTCCAAGATCCGGGAGCTCTGGNGCTGCCCATTTATGCCAACCTGCCCTCAGAC 2058
QY 430 -----434
Db 2059 ATGCAGGCTCGCATCTTCCAGCCCAACACCCCGGGGCCGGAAGAACTTCCCTCACCA 2118
QY 435 --GlnThrValAspPheSerLeuAspProThrPheThr-----446
Db 2119 TCGAAGGCATCATCTATGTCTGGACCCAGGGTTCTGCAAGCAGAGAGCTACAACCCCTC 2178
QY 447 -----447
Db 2179 GTACGGGATGAGTCACTCAACGCTCACCCCTCGCAGCAAGGCTTCAGCCCAATCAGCGGG 2238
QY 463 rgGlyArgThrGlyArgGlyLysProGlyIleTyArgPheValAlaProGlyGluArgP 483
Db 2239 CTGGCCTGCAGTCA-----483
QY 483 roSerGlyMetPheAspSerSerValLeuGlyCysGluCysTyArgPheValAlaGlyCysAlaTrp- 502
Db 2268 -----AAGTGCTCCGCTGTATACGGCTGGG 2295
QY 503 --TyrGluLeuThrProAlaGluThrThrVal 512
Db 2296 CCTATCAGCATGAGCTAGAGGAGACACAGATT 2327

RESULT 3
LOCUS CL974652
DEFINITION Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL974652
VERSION CL974652.1 GI:52403817
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
REFERENCE 1. (bases 1 to 3201)
AUTHORS Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M., Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L., Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source
1..3201
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 0.00163 Length: 3201
Score: 160.00 Matches: 80
Percent Similarity: 36.46% Conservative: 52
Best Local Similarity: 22.10% Mismatches: 128
Query Match: 4.42% Indels: 102
DB: 9 Gaps: 15

US-09-930-591-2 (1-686) x CL974652 (1-3201)
QY 183 ProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnValAlaHis 202
Db 1252 CCAATCTATAAATTCAGGGATGAACCTGCTCAAGGCTGTTGAAGAGTATCAGGTTATTGTC 1311
QY 203 LeuHisAlaProThrGlySerGlyLysSerThrLysValPro-----Ala 217
Db 1312 ATAGTAGGAGAACTCGCTCTGGTAAACACACAAATACCTCAATATCTTCATGAAGCT 1371
QY 218 AlaTyAlaAlaGlnGlyTyLysValLeuValLeuAsnProSerValAlaAlaThrMet 237
Db 1372 GGATATACAGCAAAAGCA--AAGTGGCATGTACAAACCTCGTCGAGTTGCGCAATG 1428
QY 238 GlyPheGlyValAlaTyMetSerLys-----AlaHisGlyIleAspPro 251
Db 1429 AGTGTTCAGCGAGGGTGTCCCAAGAGATGGGTGTTAACTAGGACATGAGGTTGGCTAC 1488
QY 252 AsnIleArg-----ThrGlyValArgThrIleThrThrGlySerProIleThr 267
Db 1489 TCTATAAGGTTTGAGGATTCGACGCTCTGAGAAAACATTG-----ATTAAAG 1533
QY 268 TyrSerThrTyGlyLysPheLeuAlaAsp-----GlyGlyCysSerGlyGlyAlaTy 285
Db 1534 TATATGCTGATGGCATGCTTCGAGGGAGTTCTTGGCGAACCAACATTTGGCAAGCTAT 1593
QY 286 AspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGly 305
Db 1594 AGTGTGTTATGGTTGATGAGGCTCATGAGCGTACACTGCTACTGATATCTTATTGGT 1653
QY 306 ThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeuAlaThrAlaThr 325
Db 1654 TTGGTGAAGGAT-----1665
QY 326 ProProGlySerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThr 345
Db 1666 -----ATTTCAGGTTTCACAGATCTAAAGTTGCTCATTTCAAGTGCACACC 1713
QY 346 -----GlyGluIleProPheTyGlyLysAlaIleProLeuGluAlaIleLysGly 362
Db 1714 CTTGATGCGAGAAAATTTAGTGACTACTTTGATTGAGCTCTATTTTCAAGATACCTGGG 1773
QY 363 GlyArgHisIleLeuPheCysHisSerLysLys-----LysCysAspGluLeuAlaAla 380
Db 1774 AGCGCATCTCTGTTGAAGTTTCAATATACAAAAGCTCCAGAGCAGATATATAGATGCA 1833
QY 381 LysLeuValAlaLeu-----385
Db 1834 GCCATTGTCACTGTTTTCAGATACATGTGACGCAACCCCTGGTGATATCTTGTATTC 1893
QY 386 -----GlyValAsnAlaValAlaTyTyTy-----ArgGly 395
Db 1894 CTTACAGGACAGAAAGAAATTGAACAATTGATGAATCTCTTAACACACAGAGAGGC 1953
QY 396 LeuAspValSerVal-----400
Db 1954 TTAGGCACAAAGATTGCAAGAACTACTTATCTGTCTTATATATATGCAAAATCTGCCAAGTAA 2013
QY 401 -----IleProThrSerGlyAspValValValAlaThr 412
Db 2014 CTTCAAGCTAAGATATTGAGCCACCCACAGGGGTGCTCGGAGGTTGTTCTGGCCACT 2073
QY 413 AspAlaLeuMetThrGlyPheThrGlyAsp---PheAspSerValIleAspCysAsnThr 431
```



```
QY 202 HisLeuHisAlaProThrGlySerGlyLysSerThrLysValPro----- 216
Db 228 GTGATTGTTGGAGAAACAGGATGTGGAGAGAGACCCAGATTCCACAGTACCTGGCAGAA 287
QY 217 AlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThr 236
Db 288 GCTGGCTGGACAGCGAAGACGAGTGTGGAGTGCACCGCTCGCAGAGTGGCTGCC 347
QY 237 MetGlyPheGlyAlaTyrMetSerLysAlaHisGlyLysLeuAspProAsnIleArgThrGly 256
Db 348 GATCCGATGA-TCTTTCTTCCCAAGGTGGAGCGCGAGTAGCTGATGAAGAAGGGGTGCA 406
QY 257 ValArgThrIleThrThrGlySerProIleThrTyr-----Ser 269
Db 407 GTCTGGGCCATGAAGTGGCTACTGTATCCGCTTCGATGTCACCGACCCACTGGCC 466
QY 270 ThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGly----- 282
Db 467 ACCAGAATCAAGTTCCTTACAGATGGCATGCTGTCAGAGAAATGATGTTGGATCGCGTG 526
QY 283 ---GlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIle 301
Db 527 TTAACAAATATAGTGTTCATCATCTGCGATGAAGCCACGAGAGGACCTTGTACACGGAC 586
QY 302 LeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeu 321
Db 587 ATTGCCATTGGCTGTGTGAGAGATTCAGAAAAGCGAGGGGATCTTCGTTGATTGTG 646
QY 322 AlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGlu----- 339
Db 647 GCCTCAGCCACTCGACGACGAGAAATTTTCGAGATTCTTTAAACAGAGATGAGACCACT 706
QY 340 -----ValAlaLeuSerThrThrGlyGlu----- 347
Db 707 GACCCAGCCAGAGATACCTCTGTGACGCTCACAGTGAAGAGCGGACATTTCCAGTGGAT 766
QY 348 -----IleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 767 ATCTTTTACCTACAAAGTCCTGTTCCAGATTATATCAAGCAACTGTGGACACCGTGGTA 826
QY 361 Lys-----GlyArgHisIleLeuPheCysHisSerLysLys 374
Db 827 AAAATTTCATCAGACAGAGAGATGAGACATACATAGCTTTCTTACTGCGCAGGAAGA 886
QY 375 CysAspGluLeuAlaLysLeuVal-----AlaLeuGly----- 386
Db 887 GTAGAGACTGTGTGTCATGTGATCGACAGCGCCGGCGCTGGCTCGCATCTGGGATG 946
QY 387 -----ValAsnAlaValAlaTyrTyrArgGlyLeuAsp----- 397
Db 947 AAGAAACACCTCGGGTCTCCCATGTATGACGAGTGCCTTCTTTTGACGATCAAG 1006
QY 398 ---ValSerValIleProThrSerGlyAspValValValAlaThrAspAlaLeuMet 416
Db 1007 GTGTTTGAAGGGTGTCCACAGCGCTCAGGAAGTGATCGTGCACCAATGTGGCAGAG 1066
QY 417 ThrGlyPheThr---GlyAspPheAspSerValIleAspCysAsnThrCysValThrGln 435
Db 1067 ACTTCCATCAATCACTAGTGCAATGTGTATGATTGACTGTGCTTTTATGAAGCTGCGA 1126
QY 436 ThrValAspPheSerLeuAspProThrPheThrIleGlu---ThrIleThrLeuProGln 454
Db 1127 -----GCCTACACCCAGCAGACGATTAATGAATGCTGGTGGTGGTACCAGTG 1174
QY 455 AspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGly---Ile 473
Db 1175 TCTCAGCGCTCAGCAATCAGCGGCGCAGGAGCGTGTGGCGCAACCGCTCGGGAAGTGT 1234
QY 474 TyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSerValLeuCys 493
Db 1235 TATCGCTTTACACA-----GAGGAAGCCTTTGACCGAGTACTCAGTCCACCGCTCCT 1288
QY 494 GluCysTyrAspAlaGlyCysAlaTyrTyArgLeuThrProAlaGluThrThrValArg 513
```

```
Db 1289 GAGATGACGCGCAGCAAAATTG-----GCCCCGTCATCTCGCAG 1327
QY 514 LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPheTrp 533
Db 1328 CTAAAGGCC----- 1336
QY 534 GluGlyValPheThrGlyLeuThrHisIle---AspAlaHisPheLeuSerGlnThrLys 552
Db 1337 -----CTAGGGATAGACATGCTCCTCAGGTCCACTTCATGCT----- 1375
QY 553 GlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaArgAla 572
Db 1375 ----- 1375
QY 573 GlnAlaProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLysProThr 592
Db 1376 -----CCCCCGCAGCA---CAGTCGATGGTTCAAGCCTTG----- 1408
QY 593 LeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluValThrLeu 612
Db 1409 -----CAGCTGCTCTATGCTCTCGAGGTCTGGACAAAGACTGTCGCTA 1453
QY 613 ThrHisProVal 616
Db 1454 ACTGAGCCTCTT 1465

RESULT 5
AY411078 3359 bp DNA linear GSS 16-DEC-2003
LOCUS Homo sapiens TIE gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY411078 GI:39767046
VERSION 1
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3359)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3359)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source
1..3359
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>3359
/genes="TIE"
/locus_tag="HCM4104"

gene
Alignment Scores: 0.115 Length: 3359
Pred. No.: 141.00 Matches: 168
Score: 31.38% Conservatives: 84
Percent Similarity: 20.92% Mismatches: 233
Best Local Similarity:
```

Query Match: 3.90% Indels: 318
DB: 9 Gaps: 41
US-09-930-591-2 (1-686) x AY411078 (1-3359)

QY 39 ThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyHis 58
DB 42 ACGAGCCCGGAGCGCTTCTTCTGACTTGGTGTCTGGGAGGCC----- 86
QY 59 GlyAlaGly-----ThrArg 63
DB 87 GGGCGGGGAGGGCTCGGACCGCTGGCGCGCGCGCTGCTGCTGGAGAGGACGCGT 146
QY 64 ThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAspGlnAsp--- 82
DB 147 ATCTGTGGCGACCGCGCGCGCGCGCGCTGCTGGCGCGCAACGGTTCGCACCGATC 206
QY 83 --LeuValGlyTrp-ProAlaPro----- 89
DB 207 ACGCTTCGCGCTTCTTCCAGCCCTCGGACCTCGTGGCGCTTCTCTCGTGGCGCGT 266
QY 90 -----GlnGlyAlaArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLe 107
DB 267 GCTGGCGCGCGCGCGCGCGCTATCTACGTGCACACAGCCCTGGAGCCCACTGCTT 326
QY 107 uValThrArgHis-----AlaAspValIleProVal----- 117
DB 327 CCAGACAAGTGCACACACTGTGCAACAAAGGTGACACCGCTGTACTTTCTGCACGCTG 386
QY 118 -----ArgArgArgGlyAspGlyArgGlySerLeuLeuSerProArgProI 133
DB 387 CACAAGAGAGACGACAGCGTATCTGGAAGAGCAACGGATCTACTCTACACCGCTG 446
QY 133 eSerTyrLeuLysGlySerSerGlyProLeuLeuCysProAlaGlyHisAlaValG 153
DB 447 GACTGGCATGAGCCCGAGGATGGCGGTTC-----TGTCGACCTCCCAATGTGCAG 500
QY 153 yIlePheArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPheIleProVa 173
DB 501 CCACCATCGAGCGCA-----TCTACAGTGCCTACT 530
QY 173 lGluSerLeuGluThrThrMetArgSerProValPheSerAspAsn----- 188
DB 531 TACCTGGAAGCACCGCCCTGGCGAGCGCTTCTTTCGGCTCATCGTGGCGGGTGTGGG 590
QY 189 -----SerSerProAlaValProGlnSerTyrGlnValAlaHisLeu----- 203
DB 591 GCTGGCGCTGGGGGCCAGCGCTGTACCAAGAGTGCACAGTTCCTTACATGGAGGTGTC 650
QY 204 -----HisAlaProThrGlySer---GlyLysSerThrLy 214
DB 651 TGCCACGACCATGACGGCAATGTATGCCCCCTGGCTTCACTGGCACCGCTGTGAA 710
QY 214 sValProAlaAlaTyrAla-----AlaGlnGlyTyrLy 225
DB 711 CAGGCTGTCAGAGAGGCGCGTTTGGGAGAGTGCAGGAGCAGTGCACCGGATATCA 770
QY 225 sValLeuValLeuAsnProSer-----ValAlaAlaThrMetGlyPheG 240
DB 771 GGCTGCGCGGCTTCTGCTCCAGACCCCTATGCTGCTCTTGTGGATCTGGC 830
QY 240 yAlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrI 260
DB 831 TGGAGAGGAGCCAGTGCACAGAGCTGTGCGCCCTCGTCTATTTTGGG----- 879
QY 260 eThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCy 280
DB 879 ----- 879
QY 280 sSerGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrse 300
DB 880 -----CTGATTGGCGACTCCAGTCCAGTGCAGATGTGTC 917

QY 300 rIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVa 320
DB 918 ACTTGTG---ACCGGTTCACTGGTGTGTCTGCCCCCTCTGGGTGCATGGAGTGCACGT 974
QY 320 lLeuAla---ThrAlaThrPro----- 326
DB 975 GAGAAGTCAGACCGGATCCCCCAGATCCTCAACATGGCCTCAGAACTGGAGTTCAACTTA 1034
QY 327 -----ProGlySerValThr-----ValProHisPro----- 335
DB 1035 GAGAGATGCCCCCGATCAACTGTGCAGCTGCAGGGAAACCCCTTCCCGTGGCGGAGC 1094
QY 336 -----AsnIleGluValAlaLeuSerThrGlyGluIleProPheTyrGlyLy 353
DB 1095 ATAGAGCTAGCAAGCCAGCGCACTGTCTCTCTCCACCAAGGCCATTTGTGAGCCA 1154
QY 353 sAlaIlePro-----LeuGluAlaIleLysGly 363
DB 1155 GAGAAGACCAAGCTGAGTTCGAGTCCCGCTGTTGTTCTTTCGCGACAGTGGGTCTGG 1214
QY 363 yArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaLysLeuVa 383
DB 1215 GAGTGGCGTG-----TGTCACATCTGGCGGCAAGACAGC 1250
QY 383 l-AlaLeu-GlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
DB 1251 CGGCGCTTCAAGGTCAAT-----GTGAAAGTGCCTCCC 1283
QY 403 ThrSerGlyAspValValValAlaThrAspAlaLeuMetThr----- 417
DB 1284 -----GTGCGCTGGCTGCACCTCGCTCTGTGACCAAGAGCGGCGCAG 1328
QY 418 -----GlyPheThrThrGlyAsp----- 422
DB 1329 CTTGTGTCTCCCGCTGGTCTGTTCTCTGGGATGGACCATCTCCACTGTCGCGCTG 1388
QY 423 -----PheAspSerValIleAspCysAsnThrCysValThrGlnThrValAsp 438
DB 1389 CACTACCGCGCCCGAGGACAGTACCATGACTGGTGCACCATTTGTG----- 1433
QY 439 PheSerLeuAspProThrPheThrIleGluThrIleThrLeu-----ProGln 454
DB 1434 -----GTGACCGCCAGT-----GAGAACGTGACGTTAATGAACCTGAGGCCAAAG 1478
QY 455 AspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyr 474
DB 1479 ACAGGATACAGTTCGTGTGCAGTGCAGCGCGCGCGGGAAGGA----- 1523
QY 475 ArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSer-----SerVa 491
DB 1524 -----GGAGAGGGGCTTGGGGGCTCCACCTCATGACCAAGACTGT 1568
QY 491 lLeuCysGluCysTyrAspAlaGlyCysAla---TrpTyrGluLeuThrProAlaGluThr 510
DB 1569 CCTGAGCCTTGT-----TGCAGCGTGTGTGGAGGCTGCCATGTGGAAGGC 1616
QY 511 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeu-----ProValCysGlnAsp 528
DB 1617 ACTGACCGCTGCGAGTGCAGTGGTCTTGGTCCCGGCGCCACTGGTGGGCGAC 1676
QY 529 His-----LeuGluPheTrpGluGly----- 535
DB 1677 GGTTTCTGTGCGCTGTGGGACGGGACACGGGGGCGAGCGGGGAGAACTCTCA 1736
QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeu 548
DB 1737 TCCCCCGGAGCGCGCTGCTCTCTGACGGGACTCAGCCT---GGCACCCACTACCAG 1793
QY 549 SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal 568
DB 1794 CTGGATGTGCAG-----CTCTACCACTGCACCTC 1823
QY 569 CysAlaArgAlaGlnAlaPro-----ProProSerTrpAspGlnMetTrp 583

```

Db 1824 CTGGCCCGCCTCGCCCTCGCACAGCTGCTTCTGCCCCAGTGGGCT----- 1874
QY 584 LysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeu---LeuTyrArg 602
Db 1875 -----CCAGCCCCCGACACCTCCAGCC 1898
QY 603 LeuGlyAlaValGlnAenGluValThrLeuThr-----His 614
Db 1899 CAGGCCCTCAGACTCCGAGATCCAGCTGACATGGAAGCACCAGGAGCTCTGCTGGG 1958
QY 615 ProValThrLysTyrIleMetThrCysMetSerAla----- 626
Db 1959 CCAATATCAAGTACGTGTGGAGGTGCAGGTGCTGGGGTGCAGGAGACCCACTGTGG 2018
QY 627 ---AspLeuGluValValThrSerThrTrpValLeuValGlyGlyValLeuAla 644
Db 2019 ATAGACGTGACAGCGCTGAGGAGACAGCACCATCATCGTGGCTCAACGCCAGC 2075

RESULT 6
LOCUS BU054791
DEFINITION UI-M-FDO-bzk-1-21-0-UI.r1 NIH_EMAP_FDO Mus musculus cDNA clone
IMAGE:6404900 5', mRNA sequence.
ACCESSION BU054791
VERSION BU054791.1 GI:22494868
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 822)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pyx-5.
Location/Qualifiers
1. 822
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6404900"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_EMAP_FDO"
/notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA
tail, is TGACAGAGCC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0.0146 Length: 822
Score: 140.50 Matches: 73
Percent Similarity: 35.33% Conservative: 45
Best Local Similarity: 21.86% Mismatches: 129
Query Match: 3.88% Indels: 87
DB: 5 Gaps: 11

US-09-930-591-2 (1-686) x BU054791 (1-822)
QY 98 CysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProVal 117
Db 12 TCCTGCTGCGCAGGAGGCAAGTACCAGCTGTGTGGAGGAGCAGACCATCGAGTTT 71
QY 118 ArgArgArgGlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLys 137
Db 72 GTCCGTGCTGCTCAACTCCAGGTTGAC-----GAG 101
QY 138 GlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAla 157
Db 102 GAGCCGTGAGGCCCGCCCTGTGAGCCAGGCCAGCAGAAG----- 143
QY 158 AlaValCysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGlu 177
Db 144 -----GAATCTATCCAG 155
QY 178 ThrThrMetArgSer---ProValPheSerAspAsnSerSerProProIleValProGln 196
Db 156 GCTGTGCGCGCAGGAGCTGCGCTGTTCCTTCCGAGAGAGAGCTTCTGCCCGCATTTGCC 215
QY 197 SerTyrGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValPro 216
Db 216 AACCATCAGTCTCTATCATCGAAGCGGAGAGTGGCTCTGGGAGACACACAGATCCCA 275
QY 217 -----AlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro 231
Db 276 CAGTACCTCTTTGAGGAGGGTTTACACAAAGAGGGCATGAAGATTCTTGCACCCAGCCC 335
QY 232 SerValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisGly----- 248
Db 336 CGGAGAGTGGCGGTATGAGTGTGGCAGCCCGAGTGGCGGAGAGATGGTGTGAAGCTT 395
QY 249 -----IleAspProAsnIleArg-----ThrGlyValArgThrIleThr 261
Db 396 GGAACGAGGTGGGTACAGATCCGGTTGAGGACTGCACCTCAGACGAAGTGT--- 452
QY 262 ThrGlySerProIleThrTyrSerThrTyrGly-----LysPheLeuAlaAsp 277
Db 453 -----CTCCGCTACATGACAGATGGAATGCTACTCTCCGAGAGTTCCTCTCTGAG 500
QY 278 GlyGlyCysSerGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAsp 297
Db 501 CCTGACCTTGCA-----AGTTACAGTGTGGTGAATGATGAAGAGCTCAGAGCGGACC 554
QY 298 AlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArg 317
Db 555 TTGCACACACAGATTCCTTTGGATGATCAAGAGCTCGTAGATCCGACCTCGAGCTC 614
QY 318 LeuThrValLeuAlaThrAlaThrProGlySerValThrValProHisProAsnIle 337
Db 615 AAGGTCTGTGGTTCAGCCACA----- 638
QY 338 GluGluValAlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeu 357
Db 639 -----CTGATATCTGCCGTTTCTGCTTCTTCGATGACCCCTGTCTTCA 686
QY 358 GluAlaIleLysGlyLysArgHisLeu---IlePheCysHisSerLysLysCysAspG 377
Db 687 GAATTCCTGAGCCAGCTTTCCAGTTGACATCTTCTATCAAGGCCCCAGAGGCTGACT 746
QY 377 IuLeuAlaLysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuA 397

```

```

Db      747 ACCCTGGAGCCTGGCTGGTCTGG-----CTGC 776
Qy      397 spValSerValIleProThrSerGlyAspValValVal 409
Db      777 AGATTACGCTGACCCAGCCCTGGAGATACACTGGTG 814

RESULT 7
BC059369
LOCUS
DEFINITION Homo sapiens cDNA clone IMAGE:30346256, linear HTC 08-OCT-2003
errors.
ACCESSION BC059369
VERSION BC059369.1 GI:37590735
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 3320)
Strauberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Distchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schaefer,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.J., Bosak,S.A., McEwan,P.J.,
Abramson,R.D., Mullighy,S.J., Loquellano,N.A., Peters,G.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A.C., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinska,M.I., Skalska,U., Smalhus,D.E.,
Schnierch,A., Schein,J.B., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
1247932
2 (bases 1 to 3320)
Strauberg,R.
Direct Submission
Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILLNL at: http://image.llnl.gov
Series: IRAC Plate: 133 Row: i Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 20544128
This clone has the following problem: frame shifted.
Location/Qualifiers
1..3320
/organism="Homo sapiens"

```

FEATURES
source

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30346256"
/tissue_type="placenta, normal"
/clone_id="NIH_MGC_147"
/lab_host="DH10B"
/note="Vector: pbluescript"

ORIGIN
Alignment Scores:
Pred. No.: 0.306 Length: 3320
Score: 136.50 Matches: 114
Percent Similarity: 33.69% Conservative: 73
Best Local Similarity: 20.54% Mismatches: 216
Query Match: 3.77% Indels: 153
DB: 3 Gaps: 24

US-09-930-591-2 (1-686) x BC059369 (1-3320)
Qy 132 ProfileSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAla 151
Db 35 CCGGTGAAGTTCTGGCGACCCGGTACAGAGGGCCA-----GGTGTAAAGC 79
Qy 152 ValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPheIle 171
Db 80 ATCTCTGAAGAGAGACAAAGTCTGGCTGCAAACTCTCT---GGGACAAACGGTGTATTACAAC 136
Qy 172 ProValGluSerLeu-----GluThrThrMetArgSerProValPheSerAspAsn 188
Db 137 CCTTATGCTGCCCTTTCCATAGACACAGAGGAGAGTGGCGGTATTCAAGCTTAGG 196
Qy 189 SerSerProProAlaValProGlnSerTyrGlnValAlaHisLeuHisAlaProThrGly 208
Db 197 AATCATATTTTATCTTGTAGTAAATATACAGCGGTGGTGTGTGTGTGAACACGGA 256
Qy 209 SerGlyLysSerThrLysValPro-----AlaAlaTyrAlaAlaGlnGly 223
Db 257 TGTGGGAAGAGCACACAGATTCTCAGTACCTTCGACGAAGCCGGCTGCAGACGTGAAGGA 316
Qy 224 TyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGlyAlaTyrMet 243
Db 317 AGAGTGTGTAGGAGTGACCCAGCCTCGAAGAGTGCTGCTGTACAGTTGCAGGAGAGTA 376
Qy 244 SerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIleThrThrGly 263
Db 377 GCTGAACAAGGGGTGC-AGTGTGGGCCACAGAGTGGGCTACTG-----420
Qy 264 SerProIle-----ThrTyrSerThrTyrGlyLysPheLeuAlaAspGly 278
Db 421 CATCCGCTTTGTAGTACGTCACCGACGACGCTGCCACGAGAATTAAGTTTCTTACTGTGA 480
Qy 279 GlyCysSerGly-----GlyAlaTyrAspIleIleIleCys 290
Db 481 ATGCTGCTCAGGGAATATGATGGTTGATCCGTTGTAAACAAAATATAGTGTATCATCTG 540
Qy 291 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln 310
Db 541 GATGAGCCCAACAGAGGACCTTGTACACATTCGCAATTCGCTTGTGCTGCTGCTGCTGCTG 600
Qy 311 AlaGluThrAlaGlyAlaArgLeuThrValLeuAlaThrAlaThrProGlySerVal 330
Db 601 CAGAAAAAGCGAGGGGATCTTCGATTGATTGATTAGTTCAGCCACTCTGGATGCAGACAAA 660
Qy 331 ThrValProHisProAsnIleGluGlu-----ValAla 341
Db 661 TTCCGGGATTTCTTTAATCAAAATGAACCAAGTATCCAGCAAGGGATACATGTGTGATC 720
Qy 342 LeuSerThrThrGlyGlu-----IlePro 349
Db 721 CTTACAGTGAAGGAGAACATTTCCGGTGGATATCTTTTATCTACAAAGTCTCTGTCCA 780
Qy 350 PheTyrGlyLysAlaIleProLeuGluAlaIleLys-----GlyGly 363

```


Db 781 GATTATATCAAAATCAACTGCGAAACTGTGGTGAATAATTCACACAGAGGAGCGGA 840
 Qy 364 ArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaLysLeuVal 383
 Db 841 GACGTTTACATTTCTTACTGCGCAGGAGAGGTAGAACTGTGTGTCGATGCTCATC 900
 Qy 384 -----AlaLeuGly-----ValAsnAlaValAlaLys 392
 Db 901 GACGAGGCTGACACTAGCTGCGACTGGGATGAGAGACACCTCCGAGTCTCCCCATG 960
 Qy 393 TyrArgGlyLeuAsp-----ValSerValIleProThrSerGly 405
 Db 961 TATGACGAGCTGCGCTTCTTGTAGCAAAATGAAAGTGTGTTGAAAGGTGTCACGACTGC 1020
 Qy 406 AspValValValAlaLysPheAlaLeuMetThrGlyPheThr---GlyAspPheAsp 424
 Db 1021 AGAAAGTGTATGTCGCCCAATGTCGCGCAAACTCTATCATCATCAGCGGCATTGTG 1080
 Qy 425 SerValIleAspCysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThr 444
 Db 1081 TATGTGATCGACTGTGCTTGTGAACTCCGA-----GCCATCAATCCGAG 1128
 Qy 445 PheThrIleGlu---ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArg 463
 Db 1129 ACAGCTATTGATGCTTGTGTGTGTCGCCAGTCTCCCAAGCATCAGCTAATCAGCGAGCA 1188
 Qy 464 GlyArgThrGlyArgGlyLysProGly---IleTyrArgPheValAlaProGlyGluArg 482
 Db 1189 GGACGTGTGTGCTGCTGCGGAAATGTTATCGCTTATATACA-----GAGGAA 1242
 Qy 483 ProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCysAlaLys 502
 Db 1243 GCCTTTGACAAAGTGTGCTCAGTCTACGTTCTGAGATGCGAGCTAGTAATTG----- 1296
 Qy 503 TyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThrProGly 522
 Db 1297 -----GCACCTGTCTATCTGCGTCAAGCA----- 1323
 Qy 523 LeuProValCysGlnAspHisLeuGluPheThrGlyValPheThrGlyLeuThrHis 542
 Db 1324 -----CTAGGAATGTCAAT 1338
 Qy 543 Ile---AspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeu 561
 Db 1339 GTCTCAGGTTCACCTTCATGTCG----- 1362
 Qy 562 ValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAspGln 581
 Db 1363 -----CCCCCTCCAGCA---CAGTCG 1380
 Qy 582 MetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeuTyr 601
 Db 1381 ATGGTTCAAGCCTTG-----GAGTTACTGTAT 1407
 Qy 602 ArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProVal 616
 Db 1408 GCTCTGGGAGGTCTGGACAAAGACTGTCGCTAACTGAACCGCTT 1452

RESULT 8
 CD359697
 LOCUS
 DEFINITION AGENCOURT 14284682 NIH MGC 180 Homo sapiens cDNA clone
 IMAGE:30388971 5', mRNA sequence.
 ACCESSION CD359697
 VERSION CD359697.1 GI:31131108
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 889)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgsbbs-r@mail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
<http://image.llnl.gov>
 Plate: NDAM461 row: 1 column: 04
 High quality sequence start: 22
 High quality sequence stop: 653.

FEATURES
source

Location/Qualifiers
 1..889
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30388971"
 /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
 /clone_lib="NIH_MGC_180"
 /note="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;
 Site 2: EcoRV (destroyed); Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.68 kb. Library was
 constructed by (Invitrogen). Note: this is a NIH_MGC
 Library."

ORIGIN

Alignment Scores:
 Pred. No.: 0.0771 Length: 889
 Score: 133.50 Matches: 61
 Percent Similarity: 40.24% Conservative: 40
 Best Local Similarity: 24.30% Mismatches: 95
 Query Match: 3.69% Indels: 55
 DB: 6 Gaps: 11

US-09-930-591-2 (1-686) x CD359697 (1-889)

Qy 183 ProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnValAlaHis 202
 Db 44 CCGGTGTTCCCATTTCCAGAGGAGCTCTGCTGCTATTGCAAAATCACCAGTCTCATC 103
 Qy 203 LeuHisAlaProThrGlySerGlyLysSerThrLysValPro-----Ala 217
 Db 104 ATTGAAGGCGGAGACAGCTCAGGGAAGACCCAGATCCCGCAGTATCTCTTTGAGGAG 163
 Qy 218 AlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMet 237
 Db 164 GGTATATACAAACAGGATGATGAGATGCTGCACCAACCCCGGAGAGTGGTGCATG 223
 Qy 238 GlyPheGlyAlaTyrMetSerLysAlaHisGly-----IleAspPro 251
 Db 224 AGTGTGCGCCCGAGTGTGCGCGGAGATGGGTGTGAAGCTTGGGAATGAGTTGGCTAC 283
 Qy 252 AsnIleArg-----ThrGlyValArgThrIleThrThrGlySerProIleThr 267
 Db 284 AGCATCCGTTTGGAGGACTGCATCATCAGACGCAACTGTC-----CTCCGC 328
 Qy 268 TyrSerThrTyrGly-----LysPheLeuAlaAspGlyGlyCysSerGlyGly 283
 Db 329 TACATCAGATGGGATGCTTCTCCGGGAGTCTCTCTGAGCTGACCTGGCG----- 382
 Qy 284 AlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGly 303
 Db 383 AGTTACAGCGTGTGTGTGTGGATGAGGCACACAAAGGACCCCTACACAGACATTTCTC 442
 Qy 304 IleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeuAlaThr 323
 Db 443 TTTGGATTGATCAAGGATGTTGCTCGTCCGACCTCAGCTCAAGGTCTCTGGTGGCTTCA 502

QY 324 AlaThrProGlySerValThrValProHisProAsnIleGluValAlaLeuSer 343
 DB 503 GCCACA-----ATGCACACTGCCCGTTTTC 529
 QY 344 ThrThrGlyGluLeuProPheTyrglyValAlaIleProLeuGluAlaIleTyrgly 363
 DB 530 ACC-----TTCCTTGATGACGCC-----CCTGTGTTGCAATCCCGGACGC 571
 QY 364 ArgHisLeu-----IlePheCysHisSerLysLysCysAspGluLeuAlaAlaLys 381
 DB 572 AGGTTTCCTGTGACATCTTACACCAAGGCTCCAGAGGCTGACTTGTGAAGCTTGT 631
 QY 382 LeuValAlaLeuGlyValAlaAlaValAlaTyrtyrArglyLeuAspValSerValIle 401
 DB 632 GTAGTATCTGTG-----TTGCAGATCCATGTGACC 661
 QY 402 ProThrSerGlyAspValValValAlaAlaThr 412
 DB 662 CAGCCCCCTGGGATATCCTGTGTCTTGACA 694
 RESULT 9
 LOCUS BUI187274 818 bp mRNA linear EST 04-SEP-2002
 DEFINITION AGENCOURT_7825134 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6142792
 5', mRNA sequence.
 ACCESSION BUI187274
 VERSION BUI187274.1 GI:22701258
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM13465 row: c column: 17
 High quality sequence stop: 546.
 Location/Qualifiers
 1..818
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6142792"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 67"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.0845 Length: 818
 Score: 132.50 Matches: 63
 Percent Similarity: 41.67% Conservative: 37
 Best Local Similarity: 26.25% Mismatches: 94
 Query Match: 3.66% Indels: 46
 DB: 5 Gaps: 13
 US-09-930-591-2 (1-686) x BUI187274 (1-818)

QY 128 LeuSerProArgProIleSerTyrlsLeuLysGlySerSerGlyProLeuLeuCysPro 147
 DB 50 CTCTCAGGAGCCCAAGTATCACTAGTGGTCTGAGGAGGAGGAGCACTAGTGGTTGTCCG 109
 QY 148 AlaGlyHisAlaVal-----GlyIlePheArgAlaAla-ValCysThrArgGl 163
 DB 110 ---GGCCACTCAGCTCAGGTGATGAGGAGCGCTCAGCTCCACCCTTCAACTCAGGC 166
 QY 163 yValAlaLysAlaValaAspPheIleProValGluSerLeuGluThrThrMetArgSer-- 182
 DB 167 CCAGCAGAA-----GAGTCCATCCAGGCCGTCGCCGAGCCT 205
 QY 183 -ProValPheSerAspAsnSerSerProAlaValProGlnSerTyrlsValAlaHi 202
 DB 206 CCGGTGTTCCTCCATTCGAGAGGAGCTCCTGGCTGCTATTGCAATCACCAGTCTCAT 265
 QY 202 sLeuHisAlaProThrGlySerGlyLysSerTyrlsValPro-----Al 217
 DB 266 CATTGAAGGCGAGACAGGCTCAGGGAAGACCACCAGATCCCGAGTATCTCTTTGAGGA 325
 QY 217 aAlaTyrlsAlaAlaGlnGlyTyrlsValLeuValLeuAsnProSerValAlaAlaThrMe 237
 DB 326 GGGTTATACAAACAAGGGTATGAAGATGCTGCACCCACCCGAGAGTGGCTGCCAT 385
 QY 237 tGlyPheGlyAlaTyrlsMetSerLysAlaHisGly-----IleAspPr 251
 DB 386 GAGTGTGGCCCGAGTGGCGCGGAGATGGGTGTGAAGCTTGGGAATGAGTTGGCTA 445
 QY 251 oAsnIleArg-----ThrGlyValArgThrIleThrThrGlySerProIleTh 267
 DB 446 CAGCATCCGCTTTGAGGACTGCACATCAGACGCAACTGTC-----CTCCG 490
 QY 267 rTyrlsSerTyrlsGly-----LysPheLeuAlaAspGlyGlyCysSerGlyGl 283
 DB 491 CTACATCAGACATGGGATGCTTCTCCGGAGTCTCTCTGAGGCTGACCTGGCG-- 545
 QY 283 yAlaTyrlsPheIleleleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGl 303
 DB 546 -AGTTACAGCGTGTGTGATGGTGGATGAGGCACGAAAGGACCTACACACAGACATTCT 604
 QY 303 yIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeuAlaTh 323
 DB 605 CTTTGGATTGATCAAGGATGTTGCTCGCTTCCGA---CCTGAGCTCAAGGCTCTGGTGG 661
 QY 323 xAlaThrProGlySerValThrValPro-----HisProAsnIleGluGlu 339
 DB 662 CTTCAGCCACATGGG-----ACATGCCCGGTTTTTCCACCTTCTTTTGAAGAA 713

RESULT 10
 LOCUS AL560974
 DEFINITION AL560974 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 Homo sapiens cDNA clone CS0DL006YA04 5-PRIME, mRNA sequence.
 ACCESSION AL560974
 VERSION AL560974.3 GI:46186337
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 949)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:31285103.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 18c strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a


```
Db 689 CTGCCCCACCTGTGGCACAACATCATAGAACAAAGGGAACCAACCGCGGTGTGTGT 748
Qy 18 eIeIeThrSerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSe 38
Db 749 GGAGACCAAGC-----AACCTCTCACGTGTGGCAGCAGAGGTGAC 790
Qy 38 rThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHi 58
Db 791 CTGCGCGGTGCAT-----TACCC 808
Qy 58 sGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAs 78
Db 809 CGCCCGCTGCAGAGACCAACGCCATCCAGCCTACCCCT-----ACGAG 853
Qy 78 nValAspGlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCy 98
Db 854 GACAGCATGGACTCCCGCAGCAACCCGTC-----884
Qy 98 sThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValAr 118
Db 885 -----ACCTTGCTGACCATGGACCGGCAC-----908
Qy 118 gArgArgGlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGl 138
Db 909 -----GGGGAGCAGAGCTCTATTCCCGCAGACCCCGCCTACATCCGC--953
Qy 138 ySerSerGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAl 158
Db 954 -----AGCTACCCACCCCTC-----CACCTGGACCAAGCTGGCGCTCACCGTGGCG 1003
Qy 158 aValCysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluTh 178
Db 1004 CCTGGAGCACCGGGCTACTCCCGCAGCCACCCCTTC-----1040
Qy 178 rThrMetArgSerProValPheSerAspAsnSerSerProProAla-----193
Db 1041 -----CGCAGGCCCAAGTGTAGTGGCGGCAGCTTCTCCAAGGCAGCTGCTTCTCCCA 1093
Qy 194 -----ValProGlnSerTyrGlnValAlaHisLeuHisAlaProThrGlySerGl 210
Db 1094 GTATGAGACCATGTACAGCACTACTACTTCCAGGGCTCTAGCTACCCGGAGCAGGAGG 1153
Qy 210 yLysSerThrLysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAs 230
Db 1154 GCAGTCC-----CCACCTAGCTCGCACCCCGGGGC--CGGCGCGTGCCTTCC 1201
Qy 230 nProSerValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAs 250
Db 1202 TCCGAGCGCAGTGGCAGCTGCTCTTCCCGCAGCGTGGTG-----CACGTGGCCCC 1252
Qy 250 pProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerTh 270
Db 1253 GCCTTCC-----CACCTGAGAGCGGCAGCAGCTCCAGCTTCAGCTG 1294
Qy 270 rTyr-----GlyLysPheLeuAlaAspGlyGlyCysSerGlyGl 283
Db 1295 CTATCAGCGCCACCGCTCGGTGTGAGTGGCTACCTGCGGCAC-----TGCCCGGCGAG 1348
Qy 283 y---AlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLe 302
Db 1349 CGACAGCAGCAGCAGCAGCAGCTCCGGCCAGTGGCCACTGTTCTCCAGTGACTCTGTGT 1408
Qy 302 uGlyIleGlyThrValLeuAspGln-AlaGluThrAlaGlyAla-----316
Db 1409 AGACTGCACTGAGGTGAGCAACCGAGGCGGTGTACGGGAGCTGCTCCACTTCCGCGAGCTC 1468
Qy 317 -----ArgLeuThrValLeuAlaThrAlaThrProProGlySerValThrValProHisP 335
Db 1469 CCTCAGCAGCGCACTATGACCCCTTCTATCAGCAGCGGAGCCCTCTGTCGTCAGTGA 1528
Qy 335 roAsnIleGluValAlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaI 355
Db 1529 GCGCGGGGGCTCGGGCAGCT-----CGGCGCGGGGACCTGCCTGTGCTTCGAGGGCTC 1582
```

```
Qy 355 leProLeuGluAlaIleLysGlyArgHisLeuIlePheCysHisSerLysLysLysC 375
Db 1583 CCCGCTCCCG-----AGGAGCTCCCGCGGT 1609
Qy 375 yAspGluLeuAlaAlaLysLeuValAlaLeuGlyValAsnAlaValaIleTyrTyrArgG 395
Db 1610 GCACAGTCATGCTGCTGGCGGGCGAGCCTTGGC--CGGGCCTCTCTCTCCCTCGGG 1666
Qy 395 lyLeuAspValSerValIleProThrSerGlyAspValValValAlaThrAspAlaI 415
Db 1667 GG-----1668
Qy 415 euMetThrGlyPheThrGlyAspPheAspSerValIleAspCysAsnThrCysValThrG 435
Db 1669 -----ATCAGGTGTCCACTGCAGCCTGGA 1693
Qy 435 lnThrValAspPheSerSerLeuAspProThrPheThrIleGluThrIleThrLeuProGlnA 455
Db 1694 GATGAACCTACAGCAGCAACTCTCTCCCTGGACACAGGGGGCCCAATAGCTCTACCTCAGA 1753
Qy 455 spAlaValSerArg-----ThrGlnArgArgGlyArgThrGlyArgGlyL 470
Db 1754 AGTGGGCTCAGAGCTTCTCTGGGGCGCCCTGACCTCAGAGGACCTGGAAGGGGG 1813
Qy 470 ysProGlyIleTyrArgPheValAla-----ProGlyGluA 482
Db 1814 CCACGAGTTGCGTGTGCTGCTGTCGAGGCCCGCCAGCCCTCCCGAGCGGGCCTAG 1873
Qy 482 rgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCysAlaT 502
Db 1874 CGCGGAG-----CAGCTGGCAGCAGCAGCCTTGTCTCT 1906
Qy 502 rpTyrGluLeuThrProAlaGluThrThrValArgLeuArg-----515
Db 1907 GG-----GGCCCCACCTCTACAGGGCTCTGGCCCGGGGTGGGAGCCCCAGTTC 1957
Qy 516 -----AlaTyrMet-----518
Db 1958 AGGAAGCTCCAGGGCTGTAGCGGCTTACCCCGACCATTTGCCAGGACAGATGGGT 2017
Qy 519 --AsnThrProGlyLeuProValCysGlnAspHisLeuGluPheTrpGlu-----534
Db 2018 GAAATACAGAG-GGTCTGCTGCTGCTGCTGCT-----TTCTATCAAGAGAGCAGG 2061
Qy 535 -----GlyValPheThrGlyLeuThrHisIleAspAlaH 546
Db 2062 TGGCCCGCGGGCGGAGGGGGCGAGCGGCTGCTACACT-----2099
Qy 546 isPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnA 566
Db 2100 -----GAGGACTACTCGGTGAGTGTGAGTACAGC 2130
Qy 566 laThrVal-CysAlaArgAlaGlnAlaProPro-----576
Db 2131 TCACCGAGGAACACCGCGGCTGCTACCCCGGGCCCGGACCTGAGCCAGCGCATCC 2190
Qy 577 ProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyPro 596
Db 2191 CCATCATTTCCAGAGAGTGTGAGCTGTGATCTGGGCGCTGCGCTCGGACTGCCAAGGAGCC 2250
Qy 597 ThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProVal 616
Db 2251 ACAGCTCGGCTCTGGGGTGGGA-----CGCAGGGCCGGATACCCACGGC---2298
Qy 617 ThrLysTyrIleMetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpVal 636
Db 2299 -----CCACAGCGGGCCTGGGAG 2316
Qy 637 LeuValGly-----GlyValLeuAlaAla-----LeuAlaAlaTyrCysLeuSerThr 652
Db 2317 CAACCCGGGAAGAGGAGCGGGCTCTGTGCTCCAGGCTAGGGCCCTACTGC-----2367
```

QY 653 GlyCysValValIleValGlyArgIleValIleSerGlyIleValIleProAsp 672
 Db 2368 GGCTGGCTCCCTCCGGAGAGCGGCTGTCAGGGCAACTTCCTAGTGCCTCC 2427
 QY 673 Arg 673
 Db 2428 AGG 2430

RESULT 12
 AK031534
 LOCUS
 DEFINITION Mus musculus 13 days embryo male testis cDNA, RIKEN full-length enriched library, clone:6030448F20 product:hypothetical P-loop containing nucleotide triphosphate hydrolases structure containing protein, full insert sequence.

ACCESSION AK031534
 VERSION AK031534.1 GI:26327388
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1
 Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 92279253
 PUBMED 10349636

REFERENCE 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE 4
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6
 (bases 1 to 2972)
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
 JOURNAL

COMMENT
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.jp/
 URL:http://fantom.gsc.riken.jp/.

FEATURES
 Location/Qualifiers
 1..2972
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM_DB:6030448F20"
 /db_xref="taxon:10090"
 /clone="6030448F20"
 /sex="male"
 /tissue_type="testis"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="13 days embryo"
 17..2641
 /notes="unnamed protein product; hypothetical P-loop containing nucleotide triphosphate hydrolases structure containing protein (SCOP|52540, evidence: SCOP) putative"
 /codon_start=1
 /protein_id="BAC27438.1"
 /db_xref="GI:26327389"
 /translation="MAAGRRVGPPSPRTMAPWKKRLKRLKRTGTASQGRSDSDGGEFTQABDDARAKLGPGRALPGFPPTSECVSDVEPDPTREMVRAQNKKKKSGGFSQMWL SYPVFKGIMKGVKVPPIQRTIPVLDGKDVAMARTSGTKACFLFLPFLRLKAR SAQTGARALLSPTRELALQTKFKELGKFTGLKTLALILGDKMEDQFALHNPDI IATPGRVHVAVEMNLKQSVYVVFDEADRLFMFGFAEQLEIIGRIPGGHQTULF STPLKLIYFAQAGLTPEVLRLDVSLEQLKTSFLLRVDITKAVALLYLLQNVV RPDQDQVFWATKHAELYTELIMGGQSCAHYISALDQTKARKINAKFTHNKCSSTLI VTDLAARGDIPLDNVIYSPAKGLFLHVRVARAGSTAYSLVAPDSVPYLL DLHLFGRSVTIARPCPEPSADAVGRDGVLRVQSVVDDESSLTAMGASLDLQGL LHRVANNAAQQVRSRPAAPSISIKAKELDLAELGLHPLFSSCFEGELQRLVDS LKQEEENVEVTEGVFTVEVGQRPRPGSQAKRRMETRORDQEFYVYPKPDFD SERGLSVGAGGAFEQQVAGVLDLMDGDAQNMRSQQQLKWDKRRKFRVFGSQDEK KKITBSGRFISSSYKRDLYQWKQKIDDRDSEEGPSNQRPGPRRGRGGRSQG TQSPRASSVPAGMRMRSELTKTEQILKQRRQAQKQFLQRLGGLKQLSARNRRRAQLRQ GAFGRGAPSRKGNKRM"

polyA_signal
 2954..2959
 /notes="putative"

polyA_site
 2972
 /notes="putative"

ORIGIN

Alignment Scores:
 Pred. No.: 1.34 Length: 2972
 Score: 129.00 Matches: 134
 Percent Similarity: 31.24% Conservative: 55
 Best Local Similarity: 22.15% Mismatches: 216
 Query Match: 3.57% Indels: 201
 DB: 3 Gaps: 29

US-09-930-591-2 (1-686) x AK031534 (1-2972)

Qy 100 CysGlySerSerAspLeuTyrIle-----ValThrArgHisAlaAspVal 114
 Db 90 TCGGAAACCCCACTGGGCTTCCCAAGCCGCGACACCGACTCGGATGACGGCAGT 149
 Qy 115 IleProValIargArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131

```

Db      150 TCAGATTCAGCGGAGGATGAGCGCCGGCGGAGGAGCTGGCCCTGGCAGAGCTTGC 209
Qy      132 Pro-IleSerTyrLeuLys-----GlySerSerGlyGlyProLeuLe 145
Db      210 CCTCATTCCTACCTCAGAGTCGGTATCAGATGTGGAGCCGCACACTCGGAGATGGTGC 269
Qy      145 uCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal----- 159
Db      270 GAGCCCAAGAACAGAAAAAGAAAGATCTGGAGGCTTCAGTCCATGGGCGCTGAGTTACC 329
Qy      160 -CysThr-ArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrT 179
Db      330 CTGTCTTCAAGGGGATCATGAAAAAGGCTACAGGTGCCGACGCCCATCCAGAGAGA 389
Qy      179 hrMetArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrG 199
Db      390 CCATC-----CCCGTGATCTGGATGGCAAGGATGGTGCCATG----- 430
Qy      199 InValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213
Db      431 -----GCCCGGACAGCAGTCAGTGGCAAGACGGCCTGCTTCTCTCCCGCA 473
Qy      214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn 231
Db      474 TGTTTGAGCGGCTGAAGGACGACGTGCACAGACGGGGCTCGAGCCCTCATCTCTCAC 533
Qy      231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisG 248
Db      534 CCACCGGGAGCTGGCCCTGCAGACATGAGTTCACCTAAGAGTAGGCAAGTT-CACC 592
Qy      248 ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlys 264
Db      593 GGCCTCAAGACTGCCTTGATCTGGTGGAGACAAATGGAAGACCATGTTGCAGCCCTG 652
Qy      264 erProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyA 284
Db      653 CACGAGAACCTTGACATATCATTCACCCCTCGGCGCTGCTGGTGCATGT-----GGCT 706
Qy      284 laTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyI 304
Db      707 GTGGAGATGAATGAAGCTGCAGAGTGT-----GGAGTATGTGGTG 748
Qy      304 le----- 304
Db      749 TTCGATGAAGCAGACAGGCTCTTTGAAATGGCTTTGCTGAGCAGCTACAGGATCATA 808
Qy      305 -----GlyThrValLeuAsp---GlnAlaGluThrAlaG 315
Db      809 GCGCGCCTTCTGGGGCCACAGACGGTGTCTTCTCAGCTACACTGCCCAAGCTGCTG 868
Qy      315 lyAlaArgLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332
Db      869 GTGGAATTTGCACAGCGAGGCTCAGAGCCCGTGCATCCGCTGAGCGTAGACTCC 928
Qy      332 alProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIlePro----- 349
Db      929 AAGCTCAATGACAGCTCAAGACCTC-----CTTCTCTCTGTG 967
Qy      350 -----PheTyrGlyLysAlaIleProLeuGluAlaIleLys-----GlyG 363
Db      968 GCGCAAGACACCAAGGCTGTGTCTCTCTACCTGCTGCAGATGTGGTTCGGCCCCAG 1027
Qy      363 ly-ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAlaLysLeu 382
Db      1028 GACCAGACTGGGTGTTGTAGCCCAAGAACCCATCGCGAGTAGTCTCACAGATGTGCTG 1087
Qy      383 ValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
Db      1088 ATGGCCAGGTTGAGTGGCCCACTATATAGTGCCTTGGACAGAC----- 1138
Qy      403 ThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422

```

```

Db      1139 -----GCCCGCAAGATCAACTTGGCCCAAGTTCCACACACAAC 1174
Qy      423 PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441
Db      1175 -----AAATGTTCCACCTCATCTGACTGACCTGGCCGCCGGGCGCTG 1219
Qy      442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
Db      1220 GACATCCCACTGCTGGCAACGTCATCACTACAGCTTCCCTGCCAAGGCAAGCTCTTC 1279
Qy      457 -----ValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIle 473
Db      1280 CTGACCCAGTGGCGCGTGTGGCCGACGACGCGCAAGTGGCACAGCC----- 1327
Qy      474 TyrArgPheValAlaProGlyGlu-ArgProSerSerGlyMetPheAspSerSerValLeuCy 493
Db      1328 TATCTTTGGTGGCCCCAGACGAGTCCCTTACC----- 1361
Qy      493 sGluCysTyrAspAlaGlyCysAlaTyrTyrGluLeuThrProAlaGluThrThrValAr 513
Db      1362 ---TGCTTGACCTACACCTGTTCTGGGCGCTGTGTACCTGGCCGCTCTTGTGAG 1417
Qy      513 q---LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
Db      1418 GAGCTTCAGTGGCAGATGCGTGGCAGGG-----ACGAGTGT 1456
Qy      532 eTyrGluGly----- 535
Db      1457 CTGGTCCGTCGCCAGAGTGTAGTGATGATGAGGACAGCAGCTGCAGACTCCCATG 1516
Qy      536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
Db      1517 GGGGCATCCCTGGATCTTCAGGGCTGCACC----- 1547
Qy      551 rLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571
Db      1548 -----GGTGGCCCAACAACGCTCAGCAGCAGTATGTGGCTCA 1585
Qy      571 gAlaGlnAlaPro-----ProProSerTyrAsp---GlnMetTr 583
Db      1586 CGGCAGCGCCTCGCTGAGTCCATCAAGAGAGCCAGGAGCTGGACCTGGCAGAGCTG 1645
Qy      583 pLysCys 585
Db      1646 GGCTTGC 1652

RESULT 13
AK031679
LOCUS
DEFINITION
Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
enriched library, clone:6030486C03 product:hypothetical P-loop
containing nucleotide triphosphate hydrolases structure containing
protein, full insert sequence.
ACCESSION
AK031679
VERSION
AK031679.1 GI:26327530
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374

```

```

11042159
3
PUBMED
REFERENCE
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitesunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamanoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--394-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2984)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akaihi,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-3216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
Location/Qualifiers
1. 2984
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:6030486C03"
/db_xref="taxon:10090"
/clone="6030486C03"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
25. .2754
/notes="unnamed protein product; hypothetical p-loop
containing nucleotide triphosphate hydrolases structure
containing protein (SCOP|52540, evidence: SCOP)
putative"
/codon_start=1
/protein_id="BAC27509.1"

```

/db_xref="GI:26327531"
/translation="MAAGRRVGPPSPRFTMAPWKKRLKRLKRTGTASQGRDSDSDGGE
FQIQQDDARAKLGPGRALPSPFTSECVSDVEPDTRMVRQAKKKKSGGFSQMG
SYPVFKGIMKGVKVPTELQRTIPVILGDKDVAMARTSGSGTKACFLLPMPRLKAR
SAQTGARALLISPTRELALQTKFKELGKFTGLKTALILGGDKMDQFAALHNPDI
IIATPGRLVHVAVENMLKQSVYVVDADRLHFMGPAEQLOEIIGRLPGGHQTVLF
SATPLKLVFAFAGLTPVLRLDVSINKSOLKTSFLLVRDRTKAAVLILYGLQNVV
RFQDQTVFVATKHAETLITELMGQGVSCAHYISALDQTKARKINAKITHNKSTLI
VTDLAAGLDIPLLONVINYSPAKGLFLHRVGRVARAGSGTAYSLVAPDVLQ
DLHLFLGRSVTLARPCPEPSVADAVGRDGLVRVQSVVDDSDSSITQAMGASLDLQ
LHRVANAQQQVRSPPSPSPKELDLAELGLHPLFSSCFEGELRLQRLP
IKNYRTRTTFEINAGVFTBVGGKRPSPGQAKARRMETRRDQEFVYPAPKQFD
LPQEEENVEVTEVGTBVGGKRPSPGQAKARRMETRRDQEFVYPAPKQFD
SERGLSVGAGGAFEQQVAGAVLDLMGDAQNMRSQQQLKMDKRRKRFVGGQGBDK
KKITSGSRFISSSYKDLQKWKQKIDDDSEEGFSNQRKPPRGGRGGRGQK
TQPRASSVPAGMRSELTKQILKORQSGEALPAARGPEAAFTSTPTPSPGAAP
GLWPSCSLQEGQDEKVSQTQPGWFLVAPGVDRVDYVHVLEDPSTGAAL"

polyA_signal 2963..2968
/notes="putative"
polyA_site 2984
/notes="putative"
ORIGIN

Alignment Scores:	1.35	Length:	2984
Pred. No.:	129.00	Matches:	134
Score:	31.24%	Conservative:	55
Percent Similarity:	22.15%	Mismatches:	201
Best Local Similarity:	3.57%	Gaps:	29
Query Match:	3		
DB:			

US-09-930-591-2 (1-686) x AK031679 (1-2984)

Qy	100	CysGlySerSerAspLeuTyrLeu	-----ValThrArgHisAlaAspVal	114
Db	98	TGCGGAACCCGCACTGGGGCTTCCCAAGGCGCGACACGCGACTCGATGAGCGGCAGT	157	
Qy	115	IleProValArgArg	-----GlyAspGlyArgGlySerLeuSerProArg	131
Db	158	TCGAGATCAGCGCGGAGGATGACGCGCGCGAGGAAGCTGGCGCTGGCAGACCTTGC	217	
Qy	132	Pro-IleSerTyrLeuLys	-----GlySerSerGlyGlyProLeuLe	145
Db	218	CCTCATTTCTTACCTCAGATGCGTATCAGATGCGGACCGCCACACTCGGGAGATGTGTC	277	
Qy	145	uCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal	-----	159
Db	278	GAGCCCAAGAACAGAAAGAAAGAGAGCTGGAGCTTCCAGTCCATGGGCTGAGTTACC	337	
Qy	160	-CysThr-ArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrT	179	
Db	338	CTGTGTTCAAGGGGATCATGAAAAGGGCTACAAGGTGCGCGCCCATCCAGAGGAAGA	397	
Qy	179	hrMetArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrG	199	
Db	398	CCATC-----CCCGTATCTTGGATGGCAAGGATGTGGTGGCCATG-----	438	
Qy	199	lnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr	-----	213
Db	439	-----GCCCGGACAGGCGAGTGGCAAGCGGCTGCTTCTCTCTCCGA	481	
Qy	214	-----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnP	231	
Db	482	TGTTTGAGCGGCTGAAGGCACGCGAGTGCACAGCGGGGCTCGAGCCCTCATCTCTCAC	541	
Qy	231	roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHis-	247	
Db	542	CCACCGGGAGCTGGCCCTGCAGACCATGAAGTTCAATAAGAGAGTTCAGGCAAGTT-CACC	600	
Qy	248	-----GlyIleAspProAsnIleArgThrGlyValArgThrIleThrGlyS	264	
Db	601	GGCCTCAAGACAGCCTTGATCTTGGTGGAGACAAATGGAAGACCAAGTTTGCAGCCTG	660	

QY 264 erProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyA 284
Db : : : : :
661 CAGAGAACCTGACATATCAATCAATCCACCCCTGGCGCTGTCGTGATGT-----GGCT 714
QY 284 laTyAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyI 304
Db : : : : :
715 GTGGAGATGAACCTGAAGCTGCAGAGTGT-----GGAGTATGTGTG 756
QY 304 le----- 304
Db : : : : :
757 TTCGATGAAGCAGACAGGCTCTTTGAATGGGCTTTGTGAGCAGCTACAGGAGATCATA 816
QY 305 -----GlyThrValLeuAsp---GlnAlaGluThrAlaG 315
Db 817 GGCCGCCCTTCCTGGGGCCACACAGCGGTGTCTTCTCAGCTACACTGCCAAGCTGCTG 876
QY 315 lValaArgLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332
Db 877 GTGGAAATTGCAGGGCAGGCTCACAGACCCGTCTCATCCGCTCGAGCTGAGACTGCC 936
QY 332 alProHisProAsnIleGluValAlaLeuSerThrThrGlyGluIlePro----- 349
Db 937 AAGCTCAATGAGCAGCTCAAGACCTC-----CTTCTCTCTTGTG 975
QY 350 -----PheTyGlyLysAlaIleProLeuGluAlaIleLys-----GlyG 363
Db 976 CGCGAAGACACCAAGCGTGTGTCTCTCTACCTGCGAGATGTGCTGGCGCCACG 1035
QY 363 ly-ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAlaLysLeu 382
Db : : : : :
1036 GACACAGCTGGTGTGTGACCAAGCACCATGCGGAGTACCTCCACAGATTGCTG 1095
QY 383 VallAlaLeuGlyValAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
Db : : : : :
1096 ATGGCCAGGCTGTGAGTGGCCACCATCATATAGTGTGGACACGAG----- 1146
QY 403 ThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422
Db : : : : :
1147 -----GCCCGCAAGATCAACTGGCCCAAGTTCACACACAAC 1182
QY 423 PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441
Db : : : : :
1183 -----AATGTTCCACCTCATCGTACCTGCTGCGCCGCGCGGCGCTG 1227
QY 442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
Db 1228 GACATCCACCTGTCGACACGTCATCACTACAGTTCCTCTGCCAAGGCAAGCTCTTC 1287
QY 457 -----ValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIle 473
Db 1288 CTGCACCGAGTGGCGCGTGTGGCCGAGCAGCGCCGAGTGGCAGAGCC----- 1335
QY 474 TyrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerValLeuCy 493
Db 1336 TATCTTTGGTGGCCCGCCAGAGGTCCCTTACC----- 1369
QY 493 sGluCysTyrAspAlaGlyCysAlaTyrTyrGluLeuThrProAlaGluThrValAr 513
Db 1370 ----TGCTTGACCTACACTGTCTCTGGCGCTCTGTACCTGCGCCGCTCTGTGAG 1425
QY 513 g----LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
Db 1426 GAGCTTCAGTGGCAGATGCGGTGGCAGGG-----ACCGAGTG 1464
QY 532 eTTPGluGly----- 535
Db 1465 CTGGTCCGTCGCCCCCAGAGTGTAGTGGATGATGAGGACAGACGCTGCAGACTGCCATG 1524
QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
Db 1525 GGGGCATCCCTGGATCTTCAGGGCTGACCC----- 1555
QY 551 rLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571

Db 1556 -----GGGTGGCCCAACACGCTCAGCAGCAGTATGTGGCTCA 1593
QY 571 gAlaGlnAlaPro-----ProProSerTTPAsp---GlnMetTr 583
Db 1594 CGGCAGCGCCCTCGCCTGAGTCCATCAGAGAGCCAGGAGCTGGAGCTGGCAGAGCTG 1653
QY 593 pLysCys 585
Db 1654 GGCTTGC 1660
RESULT 14
AK028274
LOCUS
DEFINITION
Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
enriched library, clone:322401E18 product:hypothetical P-loop
containing nucleotide triphosphate hydrolases structure containing
protein, full insert sequence.
ACCESSION
AK028274.1 GI:26080791
VERSION
HTC; CAP trapper.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 685-690 (2001)
MEDLINE
20530913
PUBMED
11076861
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
MEDLINE
20530913
PUBMED
11076861
6
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

538 CCACCGGGAGCTGCGCCTGCAGACCATGAATTTCACTAAAGACTAGGCAAGTT-CACC 596
||:::

248 ly-----lleAspProfsnlleAargThrGlyValAargThrlleThrThrGlys 264
:::|||||
597 GCGCTCAAGACTGCGCTTGATCTCGGTGGAGACAATAAGAACCACTTTTGAGCCGCTG 656

657 CACGAGAACCCCTGACATAATCATTCATTCGCCACCCCTGGCGCTCTGGTGCATGT-----GGCT 710

284 laTyraSpIlellellelleCysApGluCyCysHisSerThrAspAlaThrSerIleleuGlyI 304
 |||||
 |||||
 711 GTGAGATGAACCTTGAAGCTGCAGAGTGT-----GGAGTATGTGGTG 752
 304 le----- 304
 753 TTGANTGACGACAGAGGCTCTTTGAATGGGCTTTGCTGAGCAGCTACAGGAGATCAT 812

305 -----GlyTrnValLeuAsp---GlnAlaSerIleHisArg 339
||| |||
813 GGCGCGCTTCTGGGGGCCACGAGCGGTGTTCTCAGTCACATGCCCAAGCTGCTG 872
|||||

[illegible]

873 GTGGAAATTGCAAGCGCAGGCCCTCACACGAAGCCCCGCATCCCGTTCGTGCGAGTTTG
349

220 cldccu;cdproantilecgluclyvalalalaenSerThrThrglvglllepro----- 349

332 ATGTCCTCCTTGGTG 971

350 -----PheTyrGlyLysAlaIleProLeuGluAlaIleLys-----GlyC 363

972 CGCGAAGACACCAAGCTGCTGTGCTCCTCTACCTGTCAGAAATGTCGTTCGGCCCCAG 1031

363 ly-ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAlaLysLeu 382

1032 GACCAGACTGTGGTGTGTTGTAGCCACAAAGCACCATGGGAGTACCTCACAGAGTTGCTG 1091

383 ValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402

1092 ATGGGCCAGGGTGTAGTTGGCCCACTATAGTCCCTGGACCAGCG----- 1142

403 ThrSerGlyAspValValValValAlaAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422

1143 -----GCCGCGAAGATCAACTTGGCCAAAGTTCAACACACAAC 1178
 |||||
 |||||

423 PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441

1179 -----AAATGTTCCACCTCATCGTACTGACCTGGCCGCCCGGGGCGCTG 1223

442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla-----456

1224 GACATCCCACTGCTGGACAAGTCATCAACTACAGCTTCCCTGCGCAAGGCGCAAGCTCTTC 1283

```

457 -----ValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysPheGlyIle 475

```

1284 CTGCACCCAGAGGCGGCGCGTGTGGCCCCGAGCAGGCCCGAAGAGCGCACACGCC

474 TYARGPheVALALAFIOGIYGLU-HIGFOSERGRVHCHNMGSPGSGEATLRL--
||||| ||||| ||||| ||||| ::
1000 TATTTCTTCCCGCCCCAGACGGCAGGTCCCCTACC-----1365

133Z 1A11C111GG1G3CCCCGAGAGCGAGG1G3GGGGGGGG

QY 493 sGluCysTyrAspAlaGlyCysAlaTyrTyrGluLeuThrProAlaGluThrValAr 513
 Db 1366 ----TGCTGACCTACACCTGTCTCTGGCCGCTGTACACCTGGCCGCTGTGTGAG 1421
 QY 513 g---LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
 Db 1422 GAGCCTTCAGTGGCAGATCGGTTGGCAGG-----ACGGAGTG 1460
 QY 532 eTrrpGluGly----- 535
 Db 1461 CTGGGTCGCTGCCCCAGAGTGTAGTGGATGATGAGGACAGCAGCCTGCAGACTGCCATG 1520
 QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
 Db 1521 GGGGCATCCCTGGATCTTCAGGCGCTGCACC----- 1551
 QY 551 rLysGlnSerGlyGluAenLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571
 Db 1552 -----GCGTGGCCAAACAGCTCAGCAGCATGTATGTGGCTCA 1589
 QY 571 gAlaGlnAlaPro-----ProProSerTrrpAsp---GlnMetTr 583
 Db 1590 CGGCAGCGCCTCGCCTGAGTCCATCAAGAGCCCAAGAGCTGGACCTGGCAGAGCTG 1649
 QY 583 physCys 585
 Db 1650 GCCTTGC 1656

RESULT 15
 BC043699
 LOCUS BC043699 3956 bp mRNA linear HTC 13-JAN-2003
 DEFINITION Mus musculus, similar to RIKEN cDNA 2410015A15 gene, clone
 IMAGE:5344158, mRNA.
 ACCESSION BC043699
 VERSION BC043699.1 GI:27696772
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3956)
 Strausberg,R.
 Direct Submission
 Submitted (10-JAN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letitia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 86 Row: 1 Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis

This clone has the following problem: frame shifted.

FEATURES
source

Location/Qualifiers
 1. 3956
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5344158"
 /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI CGAP_Mam6"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:
 Pred. No.: 2.09 Length: 3956
 Score: 129.00 Matches: 134
 Percent Similarity: 31.24% Conservative: 55
 Best Local Similarity: 22.15% Mismatches: 216
 Query Match: 3.57% Indels: 201
 DB: 3 Gaps: 29

US-09-930-591-2 (1-686) x BC043699 (1-3956)

QY 100 CysGlySerSerAspLeuTyrLeu-----ValThrArgHisAlaAspVal 114
 Db 80 TCGGAAACGCCGAACCTGGGCTTCCCAAGCCGCGACAGCGACTCGGATGACGGCAGT 139
 QY 115 IleProValArgArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131
 Db 140 TCGAGATCCAGCGCGGAGGATGACGCCGGCGGAGAGCTGGGCCCTGGCAGACCTTGC 199
 QY 132 Pro-IleSerTyrLeuLys-----GlySerSerGlyGlyProLeuLe 145
 Db 200 CCTATTCTCTACCTCAGATGCGTATCAGATGTGGAGCCGACACTCGGAGAGTGGTGC 259
 QY 145 uCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal----- 159
 Db 260 GAGCCCAACAACAAAAAAGAGTCTGAGGCTTCCAGTCCATGGCGCTGAGTTACC 319
 QY 160 -CysThr-ArgGlyValAlaLysAlaValAppPheIleProValGluSerLeuGluThrT 179
 Db 320 CTGTGTTCAAGGGGATCATGAAAAAGGGCTACAAGGTGCCAGCCCATCCAGAGGAAGA 379
 QY 179 hrMetArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrG 199
 Db 380 CCATC-----CCCGTGATCTTGGATGGCAAGGATGGTGGCCATG----- 420
 QY 199 lnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213
 Db 421 -----GCCCGGACAGGACGAGTGGCAAGAGCGGCTGCTTCTCTCCCGA 463
 QY 214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAnp 231
 Db 464 TGTTTGACGGCTGAAGCGCAGCGATGACAGCGGGGCTCGAGCCCTCATCTCTCAC 523
 QY 231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisG 248
 Db 524 CCACCGGGAGCTGGCCCTGCAGACCATGAGTTCACTAAAGAGCTAGGCAAGTT-CACC 582
 QY 248 ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrGly 264
 Db 583 GGCCTCAAGACTGCTTGTATCTGGTGGGACAAAATGGAAGACCAAGTTTGCAGCCCTG 642
 QY 264 erProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCysSerGlyGly 284
 Db 643 CACGAGAACCTGACATAATCATTTGCCACCCCTCGGCGCTCTGGTGCATGT-----GGCT 696
 QY 284 laTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyI 304
 Db 697 GTGGAGATGAACCTTGAAGCTGCAGAGTGT-----GGAGTATGTGTGTG 738
 QY 304 le----- 304

Mon Feb 28 09:44:43 2005

us-09-930-591-2.rst

```
Db      739  TTCCGATGAGCAGACAGAGGCTCTTTGAAATGGGCTTTGCTGAGCAGCTACAGAGATCATA 798
Qy      305  -----GlyThrValLeuAsp---GlnAlaGluThrAlaG 315
Db      799  GGGCGCTTCTCGGGGCCACAGAGGGTCTGTCTCAGCTACACTGCCCAAGCTGCTG 858
Qy      315  lyAlaArgLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332
Db      859  GTGGAATTTGACGGCGAGGCTCAGAGAGCCGCTCATCGCTGAGCTGAGACTCC 918
Qy      332  alProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIlePro----- 349
Db      919  AAGCTCAATGAGCAGCTCAAGACCTC-----CTTCCTCTCTGTG 957
Qy      350  -----PheTyrGlyLeuAlaIleProLeuGluAlaIleLys-----GlyG 363
Db      958  CGCGAAGACACCAAGGCTGTGCTCTCTACTGCTGAGAATGTCGTTCGGCCCCAG 1017
Qy      363  ly-ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaLysLeu 382
Db      1018  GACCAGACTGTGTGTTGTAGCCACAAGCACCATCGGAGTACCTCAGAGATGCTG 1077
Qy      383  ValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
Db      1078  ATGGCCAGGCTGTGAGTGGCCCCACATCTATAGTCCTTGGACACAGC----- 1128
Qy      403  ThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422
Db      1129  -----GCCGCAAGATCAACTTGGCCAAAGTTTCACACACAAC 1164
Qy      423  PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441
Db      1165  -----AAATGTTCCACCTCATCGTACTGACCTGGCGCGCCGGGCTG 1209
Qy      442  AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
Db      1210  GACATCCCACTGCTGACACAGTCACTCACTCCCTGCTGCCAAGGGCAAGCTCTTC 1269
Qy      457  -----ValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyLe 473
Db      1270  CTGCACCGAGTGGCGCGTGTGGCCGAGCAGCGCGAAGTGGCAGCC----- 1317
Qy      474  TyrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerSerValLeuCy 493
Db      1318  TATTCTTGTGGCCCGCAGACAGAGTCCCTACC----- 1351
Qy      493  sGluCysTyrAspAlaGlyCysAlaTyrTyrGluLeuThrProAlaGluThrValAr 513
Db      1352  -----TGCCTGACCTACACTGTTCTCGGCGCGCTGTACCCCTGGCCCGCTCTGTGAG 1407
Qy      513  g---LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
Db      1408  GAGCCTTCAGTGCAGATCGGTGGCAGG-----ACGAGTG 1446
Qy      532  eTrpGluGly----- 535
Db      1447  CTGGGTGCGGTCGCCCGAGTGTAGTGGATGATGAGCAGCAGCCTGCAGACTGCCATG 1506
Qy      536  -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
Db      1507  GGGGCATCCCTGGATCTTCAGGCGCTGCACC----- 1537
Qy      551  rLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571
Db      1538  -----CCGTGGCCACAACGCTCAGCAGCAGTATGTGGCTCA 1575
Qy      571  gAlaGlnAlaPro-----ProProSerTrpAsp---GlnMetTr 583
Db      1576  CGGCCAGCGCCTCGCTGAGTCCATCAAGAGAGCCAGGAGCTGGACCTGGCAGAGCTG 1635
Qy      583  pLysCys 585
```

Db 1636 GGCTTGC 1642

Search completed: February 27, 2005, 07:18:59
Job time : 4768 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 15:23:02 ; Search time 8681 Seconds
(without alignments)
11503.997 Million cell updates/sec

Title: US-09-930-591-1
Perfect score: 2061
Sequence: 1 atggcgccctacgcgcta.....atgaatgaagagtgtga 2061

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	Length	DB	ID	Description
1	2061	100.0	2061	6	CQ826998		CQ826998 Sequence
2	2061	100.0	2061	6	AX441176		AX441176 Sequence
3	2061	100.0	2061	6	AX467113		AX467113 Sequence
4	1854.8	90.0	9610	14	HEC278830		HEC278830 Sequence
5	1846.4	89.6	8791	14	AY615798		AY615798 Sequence
6	1787	86.7	6299	6	AX164584		AX164584 Sequence
7	1786	86.7	5360	6	I06434		I06434 Sequence
8	1786	86.7	5360	6	I09328		I09328 Sequence
9	1786	86.7	6785	6	I06440		I06440 Sequence
10	1786	86.7	7310	6	AR118696		AR118696 Sequence
11	1786	86.7	7310	6	I09331		I09331 Sequence
12	1786	86.7	7310	14	HPCPOLYP		HPCPOLYP Sequence
13	1786	86.7	9185	6	I08294		I08294 Sequence
14	1786	86.7	9185	6	BD091382		BD091382 Sequence
15	1786	86.7	9379	6	AR166930		AR166930 Sequence
16	1786	86.7	9379	6	AR301300		AR301300 Sequence
17	1786	86.7	9401	6	AR176483		AR176483 Sequence
18	1786	86.7	9401	6	E66593		E66593 Sequence
19	1786	86.7	9401	6	I71894		I71894 Sequence

20	1786	86.7	9401	6	I81885		I81885 Sequence
21	1786	86.7	9401	6	BD080334		BD080334 Sequence
22	1786	86.7	9401	14	HPCPOLYP		HPCPOLYP Sequence
23	1784.4	86.6	6785	6	AR118692		AR118692 Sequence
24	1784.4	86.6	6785	6	I09329		I09329 Sequence
25	1784.4	86.6	8316	6	AR118703		AR118703 Sequence
26	1784.4	86.6	8987	6	AR118728		AR118728 Sequence
27	1784.4	86.6	9185	6	AR118722		AR118722 Sequence
28	1784.4	86.6	9185	6	AR118723		AR118723 Sequence
29	1784.4	86.6	9379	6	AR118747		AR118747 Sequence
30	1781.2	86.4	2058	6	AR404933		AR404933 Sequence
31	1781.2	86.4	2058	6	AR395309		AR395309 Sequence
32	1781.2	86.4	5360	6	AR118686		AR118686 Sequence
33	1781.2	86.4	9424	14	AF511948		AF511948 Sequence
34	1781.2	86.4	9609	12	AF387805		AF387805 Sequence
35	1781.2	86.4	9618	14	AF271632		AF271632 Sequence
36	1781.2	86.4	9646	12	AF387806		AF387806 Sequence
37	1781.2	86.4	9693	12	AF387807		AF387807 Sequence
38	1779.6	86.3	2058	6	AR408362		AR408362 Sequence
39	1779.6	86.3	2058	6	AX454818		AX454818 Sequence
40	1774.8	86.1	12980	6	AR110831		AR110831 Sequence
41	1774.8	86.1	12980	6	BD069985		BD069985 Sequence
42	1773.2	86.0	9609	12	AF387808		AF387808 Sequence
43	1773.2	86.0	9646	6	AR110828		AR110828 Sequence
44	1773.2	86.0	9646	6	BD069982		BD069982 Sequence
45	1773.2	86.0	9646	14	AF009606		AF009606 Sequence

ALIGNMENTS

RESULT 1
CQ826998
LOCUS CQ826998 2061 bp DNA linear PAT 29-JUN-2004
DEFINITION Sequence 1 from Patent WO2004048402.
ACCESSION CQ826998
VERSION CQ826998.1 GI:49455655
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Sallberg, M.
TITLE A hepatitis c virus codon optimized non-structural ns3/4a fusion gene
JOURNAL Patent: WO 2004048402-A 1 10-JUN-2004;
TRIPEP AB (SE)
FEATURES
source
1. .2061
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Hepatitis C virus NS3/4A coding region"

ORIGIN

Query Match 100.0%; Score 2061; DB 6; Length 2061;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCCTTATCAGCGCCTATGCCAGCAGACAGAGGGGCTTTTGGGATGCATAATCACC 60
Db 1 ATGGCGCCTTATCAGCGCCTATGCCAGCAGACAGAGGGGCTTTTGGGATGCATAATCACC 60
QY 61 AGCTTGACCGCGCGGACAAAACACGAGTGGAGGTTCAGATCGTGCACTGCT 120
Db 61 AGCTTGACCGCGCGGACAAAACACGAGTGGAGGTTCAGATCGTGCACTGCT 120
QY 121 GCCCAGACTTCTTGGCACTGCATTAACGGGGTGTGTGACTCTCTACCATGAGCC 180
Db 121 GCCCAGACTTCTTGGCACTGCATTAACGGGGTGTGTGACTCTCTACCATGAGCC 180
QY 181 GGAACAAGGACCATTCGTCCTCAAGGGTCTCTTATCCAGATGTACCAATGTGGAC 240

181 GGAAACAAGGACCAATTGCGTCACTAAGGGTCTCTGTTATCCAGATGTACACCAATGTGGAC 240
241 CAAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCGCGCTCATTAACACATGCACTTGC 300
241 CAAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCGCGCTCATTAACACATGCACTTGC 300
301 GGTCTCTCGGACCTTTACCTGTGTGACAGGACCGCGATGTCAATCTCTGTGCGCGAGCG 360
301 GGTCTCTCGGACCTTTACCTGTGTGACAGGACCGCGATGTCAATCTCTGTGCGCGAGCG 360
361 GGTGATGGCAGGGGACGCTGCTTTGCGCCCGGCTTATCTTACTTTGAAGGCTCTCG 420
361 GGTGATGGCAGGGGACGCTGCTTTGCGCCCGGCTTATCTTACTTTGAAGGCTCTCG 420
421 GGAGGCCCTCTGCTGTGCGCCCGCAGGACATGCGGTAGGCATATTAGAGCCCGGTATGC 480
421 GGAGGCCCTCTGCTGTGCGCCCGCAGGACATGCGGTAGGCATATTAGAGCCCGGTATGC 480
481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGCTTTAGAGACAACCATG 540
481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGCTTTAGAGACAACCATG 540
541 AGGTCCCGGCTGTCTCAGACAATCTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG 600
541 AGGTCCCGGCTGTCTCAGACAATCTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG 600
601 GCCACCTCGTCTCCACCGCAGCGGTAGAGACCAAGTCCCGCCCGCATACGCA 660
601 GCCACCTCGTCTCCACCGCAGCGGTAGAGACCAAGTCCCGCCCGCATACGCA 660
661 GCTCAGGGCTACAAGGTGCTGTGCTCAACCCCTCGTTGCTGCAACAATGGCTTTGCT 720
661 GCTCAGGGCTACAAGGTGCTGTGCTCAACCCCTCGTTGCTGCAACAATGGCTTTGCT 720
721 GCTTACATGTCCAAGGCCCATGGGATTTGATCTTAACATCAGGACTGGGGTGAGGCAATT 780
721 GCTTACATGTCCAAGGCCCATGGGATTTGATCTTAACATCAGGACTGGGGTGAGGCAATT 780
781 ACTACTGGCAGCCGATCAGTATTCACCTAGGCAAGTCTTCTGCGACCGGGGT 840
781 ACTACTGGCAGCCGATCAGTATTCACCTAGGCAAGTCTTCTGCGACCGGGGT 840
841 TCAGGGGTGCTTATGACATAATAATTTGTGACAGTGCCACTCCACCGATGCAACATCC 900
841 TCAGGGGTGCTTATGACATAATAATTTGTGACAGTGCCACTCCACCGATGCAACATCC 900
901 ATCTTGGGCAATGGCACTGCTCTTGACCAAGCAGAGACCGCGGGGCGAGCTGACTGTG 960
901 ATCTTGGGCAATGGCACTGCTCTTGACCAAGCAGAGACCGCGGGGCGAGCTGACTGTG 960
961 CTGCGCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCCCATCTTAACATCGAGGATT 1020
961 CTGCGCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCCCATCTTAACATCGAGGATT 1020
1021 GCTCTGTCCACTACCGGAGATCCCTTTTATGGCAAGGCTATTCCTCTGAAGCAATT 1080
1021 GCTCTGTCCACTACCGGAGATCCCTTTTATGGCAAGGCTATTCCTCTGAAGCAATT 1080
1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTGCGACGAGCTCGCCGCA 1140
1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTGCGACGAGCTCGCCGCA 1140
1141 AAATGGTTCGGTGGGCTCAATGCGGTGGCTTACTACCGGGCTTATGATGTCCTGTC 1200
1141 AAATGGTTCGGTGGGCTCAATGCGGTGGCTTACTACCGGGCTTATGATGTCCTGTC 1200
1201 ATCCCCCAGGAGTGTGACGTTGCTGCGTGGCAATGACGCCCTCATGACCGGCTTACC 1260
1201 ATCCCCCAGGAGTGTGACGTTGCTGCTGCGTGGCAATGACGCCCTCATGACCGGCTTACC 1260
1261 GCGCACTTCGATTCGGTGTAGACTGCAACAGGTGTGTCACCGAGACAGTCTGACCTCAGC 1320
1261 GCGCACTTCGATTCGGTGTAGACTGCAACAGGTGTGTCACCGAGACAGTCTGACCTCAGC 1320

1321 CTTGACCTTACCTTCAACCATTTGAGACAATCACGTTCCCGAGATGCTCTCTCCGTACT 1380
1321 CTTGACCTTACCTTCAACCATTTGAGACAATCACGTTCCCGAGATGCTCTCTCCGTACT 1380
1381 CAACGTCTGGGTAGGACTGGCAGAGGGAAGCCAGGCATCTACAGATTTTGTGGCACCGGG 1440
1381 CAACGTCTGGGTAGGACTGGCAGAGGGAAGCCAGGCATCTACAGATTTTGTGGCACCGGG 1440
1441 GAGCGTCTTCTGCGATGTTTGAATCTGCTCTCTCTGCGAGTCTATGACGCGGTCT 1500
1441 GAGCGTCTTCTGCGATGTTTGAATCTGCTCTCTCTGCGAGTCTATGACGCGGTCT 1500
1501 GCTTGGTATGAGCTTACGCCCGCCGAGACCAAGTTAGCTTACGAGCATACATGACACC 1560
1501 GCTTGGTATGAGCTTACGCCCGCCGAGACCAAGTTAGCTTACGAGCATACATGACACC 1560
1561 CCGGACTTCCCGTGTGCAAGACCATCTTGAATTTTGGGAGGGCTCTTTACGGGTCTC 1620
1561 CCGGACTTCCCGTGTGCAAGACCATCTTGAATTTTGGGAGGGCTCTTTACGGGTCTC 1620
1621 ACCCATAGAGCCCACTTCTATCCAGACAAGAGAGTGGGGAACCTTCCCTAT 1680
1621 ACCCATAGAGCCCACTTCTATCCAGACAAGAGAGTGGGGAACCTTCCCTAT 1680
1681 CTGTAGCTTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGCTGGAC 1740
1681 CTGTAGCTTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGCTGGAC 1740
1741 CAGATGTGGAAGTGTGATCCGCTCAAGCCCACTCCATGGGCCAACCTCTGCTA 1800
1741 CAGATGTGGAAGTGTGATCCGCTCAAGCCCACTCCATGGGCCAACCTCTGCTA 1800
1801 TATAGACTGGGCGCTGTCCAGAAATGAAGTCACTGACGACCCAGTCAACCAAGTATATC 1860
1801 TATAGACTGGGCGCTGTCCAGAAATGAAGTCACTGACGACCCAGTCAACCAAGTATATC 1860
1861 ATGACATGTATGTGCGCTGACCTGAGGTGCTGACAGTACCTGAGTCTGCTGCGGC 1920
1861 ATGACATGTATGTGCGCTGACCTGAGGTGCTGACAGTACCTGAGTCTGCTGCGGC 1920
1921 GTTCTGGCTGCTTTGGCCGCTATTGCTTATCCATCAGGCTGCTGCTCATAGTAGG 1980
1921 GTTCTGGCTGCTTTGGCCGCTATTGCTTATCCACAGGCTGCTGCTCATAGTAGG 1980
1981 ATTGCTTGTCCGGAAGCCGCAATCATACCGACAGGAAGTCTCTACCGGAGTTC 2040
1981 ATTGCTTGTCCGGAAGCCGCAATCATACCGACAGGAAGTCTCTACCGGAGTTC 2040
2041 GATGAATGGAAGTGTCTGA 2061
2041 GATGAATGGAAGTGTCTGA 2061

RESULT 2

AX441176
LOCUS AX441176 2061 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 16 from Patent WO0213855.
ACCESSION AX441176
VERSION AX441176.1 GI:21665758
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Sallberg, M. and Hultgren, C.
TITLE Vaccines containing ribavirin and methods of use thereof
JOURNAL Patent: WO 0213855-A 16 21-FEB-2002;
TRIPEP AB (SE)
FEATURES
Location/Qualifiers
1..2061
/organism="synthetic construct"
/mol_type="unassigned DNA"

/db_xref="taxon:32630" /note="Hepatitis C virus NS3/4A coding region"									
ORIGIN									
Query Match	100.0%;	Score	2061;	DB	6;	Length	2061;		
Best Local Similarity	100.0%;	Pred. No.	0;	Mismatches	0;	Indels	0;	Gaps	0;
Matches	2061;	Conservative	0;						
QY	1	ATGGGCGCTATCAGCGCTATGCCCAGACAGACAAGGGGCTTTTGGGATGCATAATCACC	60						
DB	1	ATGGGCGCTATCAGCGCTATGCCCAGACAGACAAGGGGCTTTTGGGATGCATAATCACC	60						
QY	61	AGCTTGACCGCGGAGCAAAAACACAGGTGAGAGGTGAGGTTCAGATCGTGTCAACTGCT	120						
DB	61	AGCTTGACCGCGGAGCAAAAACACAGGTGAGAGGTGAGGTTCAGATCGTGTCAACTGCT	120						
QY	121	GCCAGACTTTCTGGCAACTGCATTAACGGGGTGTGGACTGTCTACCATGGAGCC	180						
DB	121	GCCAGACTTTCTGGCAACTGCATTAACGGGGTGTGGACTGTCTACCATGGAGCC	180						
QY	181	GGAAACAGGACCAATTCGCTCACTAAGGGTCTCTGTTATCCAGATGTACACCAATGTGGAC	240						
DB	181	GGAAACAGGACCAATTCGCTCACTAAGGGTCTCTGTTATCCAGATGTACACCAATGTGGAC	240						
QY	241	CAAGACCTCGTAGGGTGGCCCGCTCCCAAGGTGCCCGTCAATTAACCATGCACTTGC	300						
DB	241	CAAGACCTCGTAGGGTGGCCCGCTCCCAAGGTGCCCGTCAATTAACCATGCACTTGC	300						
QY	301	GGCTCTCGGACTTTACCTGTGTGACGAGGACCGCGATGTCATTCCTGTGCGCGACGG	360						
DB	301	GGCTCTCGGACTTTACCTGTGTGACGAGGACCGCGATGTCATTCCTGTGCGCGACGG	360						
QY	361	GGTGATGGCAGGGCAGCTGCTTTCCGCCCGGCTCTCTCTTACTTTGAAAGGCTCCTCG	420						
DB	361	GGTGATGGCAGGGCAGCTGCTTTCCGCCCGGCTCTCTCTTACTTTGAAAGGCTCCTCG	420						
QY	421	GGAGGCGCTCTGTGTGCGCGCAGACATGCGGTAGGCATATTCAGAGCGCGGTATGC	480						
DB	421	GGAGGCGCTCTGTGTGCGCGCAGACATGCGGTAGGCATATTCAGAGCGCGGTATGC	480						
QY	481	ACCCGTGGAGTGGCTAAGCGGTGAGCTTCATCCCGGTAGAGGCTTAGAGCAACCATG	540						
DB	481	ACCCGTGGAGTGGCTAAGCGGTGAGCTTCATCCCGGTAGAGGCTTAGAGCAACCATG	540						
QY	541	AGGTCCCGGCTTCTCAGACAACTCCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG	600						
DB	541	AGGTCCCGGCTTCTCAGACAACTCCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG	600						
QY	601	GCCACCTGCTATGCTCCCAACCGCAGCGGTAAAGACCAAGGTCCCGCGCGATACGCA	660						
DB	601	GCCACCTGCTATGCTCCCAACCGCAGCGGTAAAGACCAAGGTCCCGCGCGATACGCA	660						
QY	661	GCTCAGGGCTACAGGTGTGGTGTCAACCCCTCCGTTGCTGCAACATGGGCTTGGT	720						
DB	661	GCTCAGGGCTACAGGTGTGGTGTCAACCCCTCCGTTGCTGCAACATGGGCTTGGT	720						
QY	721	GCTTACATGTCCAAAGGCCCATGGATTCCTTAACATCAGGACTGGGCTGAGGACAAAT	780						
DB	721	GCTTACATGTCCAAAGGCCCATGGATTCCTTAACATCAGGACTGGGCTGAGGACAAAT	780						
QY	781	ACTACTGGCAGCGCGATCACTATCCACTACGGCAAGTTCCTCCCAACCGCGGTGT	840						
DB	781	ACTACTGGCAGCGCGATCACTATCCACTACGGCAAGTTCCTCCCAACCGCGGTGT	840						
QY	841	TCAGGGGGTCTTATGACATAATAATTTGTGACGAGTGCCTCCAGCATGGATCCCAATCC	900						
DB	841	TCAGGGGGTCTTATGACATAATAATTTGTGACGAGTGCCTCCAGCATGGATCCCAATCC	900						
QY	901	ATCTTGGGCAATTCGCACTGTCTTGTACCAAGCAGACACCGCGGGCGAGACTGTG	960						
DB	901	ATCTTGGGCAATTCGCACTGTCTTGTACCAAGCAGACACCGCGGGCGAGACTGTG	960						
QY	961	CTCGCCACCGCTACCCCTCCGGGCTCCGTCACGTGTGCCCCATCTTAACATCGAGGAGTT	1020						
DB	961	CTCGCCACCGCTACCCCTCCGGGCTCCGTCACGTGTGCCCCATCTTAACATCGAGGAGTT	1020						
QY	1021	GCTCTGTCCACTACCGGAGAGATCCCTTTTATGTGCAAGGCTATTCCCTTGAAGCAAT	1080						
DB	1021	GCTCTGTCCACTACCGGAGAGATCCCTTTTATGTGCAAGGCTATTCCCTTGAAGCAAT	1080						
QY	1081	AAGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAAAGTGCGAGCTCGCGCA	1140						
DB	1081	AAGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAAAGTGCGAGCTCGCGCA	1140						
QY	1141	AAACTGTGTGGGCTCAATCGCGTGTACTACCGGGCTTCAATGTGTCCGTC	1200						
DB	1141	AAACTGTGTGGGCTCAATCGCGTGTACTACCGGGCTTCAATGTGTCCGTC	1200						
QY	1201	ATCCCGACAGTGTGTGCTGTCTGTGGCAACTGACCGCCCTCATGACCGGCTTTACC	1260						
DB	1201	ATCCCGACAGTGTGTGCTGTCTGTGGCAACTGACCGCCCTCATGACCGGCTTTACC	1260						
QY	1261	GGCGACTTTCGATTCGATGACATGCAACAGTGTGTCCAGGATGCTCTCCGCTACT	1320						
DB	1261	GGCGACTTTCGATTCGATGACATGCAACAGTGTGTCCAGGATGCTCTCCGCTACT	1320						
QY	1321	CTTGACCTTACCTTCAACATTTGAGACAAATCACGTTTCCCGAGATGCTGTCCCGTACT	1380						
DB	1321	CTTGACCTTACCTTCAACATTTGAGACAAATCACGTTTCCCGAGATGCTGTCCCGTACT	1380						
QY	1381	CAACGTCTGGGTGAGGCTGGCAGAGGAGGAGCCAGGCATCTACAGATTTTGTGGCACCGGG	1440						
DB	1381	CAACGTCTGGGTGAGGCTGGCAGAGGAGGAGCCAGGCATCTACAGATTTTGTGGCACCGGG	1440						
QY	1441	GAGCGTCTTCTGGCATGTTGACTGCTGTCTCTCTGCGAGTGTATACGCGGCTTGT	1500						
DB	1441	GAGCGTCTTCTGGCATGTTGACTGCTGTCTCTCTGCGAGTGTATACGCGGCTTGT	1500						
QY	1501	GCTTGTGTAGCTTACGCGCGCGCAGACCAAGTTAGCTTACGAGCATACATGACACC	1560						
DB	1501	GCTTGTGTAGCTTACGCGCGCGCAGACCAAGTTAGCTTACGAGCATACATGACACC	1560						
QY	1561	CCGGGACTTCCGCTGTGCGCAAGACCATCTTGAATTTTGGAGGGGCTCTTTACGGGTCTC	1620						
DB	1561	CCGGGACTTCCGCTGTGCGCAAGACCATCTTGAATTTTGGAGGGGCTCTTTACGGGTCTC	1620						
QY	1621	ACCCACATAGACGCCACTTCTATCCAGACAAAGAGAGTGGGAAAACTTCCCTAT	1680						
DB	1621	ACCCACATAGACGCCACTTCTATCCAGACAAAGAGAGTGGGAAAACTTCCCTAT	1680						
QY	1681	CTGTGTAGCTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGGCTGGGAC	1740						
DB	1681	CTGTGTAGCTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGGCTGGGAC	1740						
QY	1741	CAGATGTGGAAGTGTGATCCGCTCTCAAGCCACCCCTCATGGGCGCAACACTCTGTCTA	1800						
DB	1741	CAGATGTGGAAGTGTGATCCGCTCTCAAGCCACCCCTCATGGGCGCAACACTCTGTCTA	1800						
QY	1801	TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATATC	1860						
DB	1801	TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATATC	1860						
QY	1861	ATGACATGTATGTGGCTGACCTGAGGTGCTCAAGTACCTGGGTGCTCTGTGGGCGC	1920						
DB	1861	ATGACATGTATGTGGCTGACCTGAGGTGCTCAAGTACCTGGGTGCTCTGTGGGCGC	1920						
QY	1921	GTTCTGGCTGCTTGGCGCGTATTGCTTATCCAGAGGTGCTGTGTGTATAGTAGTAGG	1980						
DB	1921	GTTCTGGCTGCTTGGCGCGTATTGCTTATCCAGAGGTGCTGTGTGTATAGTAGTAGG	1980						
QY	1981	ATTGTCTTGTCCGAAAGCCGCAATCATACCCGACAGGAAAGTCTCTTACCGGGAGTTC	2040						
DB	1981	ATTGTCTTGTCCGAAAGCCGCAATCATACCCGACAGGAAAGTCTCTTACCGGGAGTTC	2040						
QY	2041	GATGAAATGGAAGTGTCTGA	2061						


```
QY 1741 CAGATGTGAAGTGTGATTCCTCTCAAGCCCAACCTCCATGGGCAACACCTCTGCTA 1800
Db 1741 CAGATGTGAAGTGTGATTCCTCTCAAGCCCAACCTCCATGGGCAACACCTCTGCTA 1800
QY 1801 TATAGACTGGCGCTCTCAGATGAAGTCAACCTGACGACCCAGCTACCAAGTATATC 1860
Db 1801 TATAGACTGGCGCTCTCAGATGAAGTCAACCTGACGACCCAGCTACCAAGTATATC 1860
QY 1861 ATGACATGTATGTGCGCTGACCTGAGAGTCTGTCACGAGTACCTGGGTGCTCGTTGGCGGC 1920
Db 1861 ATGACATGTATGTGCGCTGACCTGAGAGTCTGTCACGAGTACCTGGGTGCTCGTTGGCGGC 1920
QY 1921 GTTCTGGCTGCTTTGGCGGTAATGCTATCCTATCACAAGGCTGCTGCTATAGTAGTAGG 1980
Db 1921 GTTCTGGCTGCTTTGGCGGTAATGCTATCCTATCACAAGGCTGCTGCTATAGTAGTAGG 1980
QY 1981 ATTGTCTGTCGCGAAGCGGCAATCATACCCGACGAGGAGTCTCTACCGGAGTTC 2040
Db 1981 ATTGTCTGTCGCGAAGCGGCAATCATACCCGACGAGGAGTCTCTACCGGAGTTC 2040
QY 2041 GATGAATGGAAGTGTCTGA 2061
Db 2041 GATGAATGGAAGTGTCTGA 2061

;
;
;
LOCUS HEC278830 9610 bp RNA linear VRL 03-JAN-2001
DEFINITION Hepatitis C virus genomic RNA for polyprotein gene.
ACCESSION AJ278830
VERSION AJ278830.1 GI:9843676
KEYWORDS core protein; envelop protein 1; envelop protein 2; non-structural protein 2; non-structural protein 3; non-structural protein 4a; non-structural protein 4b; non-structural protein 5A; non-structural protein 5B; ORF1; ORF10; ORF2; ORF3; ORF4; ORF5; ORF6; ORF7; ORF8; ORF9; polyprotein.
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepadnavirus.
REFERENCE 1
AUTHORS Kumar, U., Tuthill, T., Thomas, H.C. and Monjardino, J.
TITLE Sequence, expression and reconstitution of an HCV genome from a British isolate derived from a single blood donation
JOURNAL J. Viral Hepat. 7 (6), 459-465 (2000)
MEDLINE 21014672
PUBMED 11115058
REFERENCE 2 (bases 1 to 9610)
AUTHORS Kumar, U.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-2000) Kumar U., Virology, GlaxoWellcome Research centre, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2NY, UNITED KINGDOM
FEATURES
source Location/Qualifiers
1 .9610
/organism="Hepatitis C virus"
/virus
/mol_type="genomic RNA"
/db_xref="taxon:11103"
/notes="derived from serum of a non-A, non-B hepatitis patient"
342..9377
/notes="unnamed protein product; ORF1"
/codon_start=1
/protein_id="CAC03609.1"
/db_xref="GI:9843677"
/db_xref="GOA:Q9DIT6"
/db_xref="UniProt/TsEMBL:Q9DIT6"
/translators="MSTNPKPQRTKENTNRRPQDVKFGGGOI VGGVLLPRGPRRL
GVATRKTSRSRPRGRQPIKARPEGTWAOCPGPLYGNECGWAGWLLSPRG
SRPSWGTPDPRRNRKVIDITLCGFADLMGIYIPLVGFLGGAARALAHGVRVLED
GVNATGNLPCGFSFLLALLSCLTPASAYHVRNSGLYHVTNCPNNSIYVETAD
TILHSPGCVPCVREGNASRCWAVPTVATRDGKLPPTQLRRHIDLLVGSATLCSALY
```

mat_peptide

/product="core protein"

/notes="ORF2"

915..1490

mat_peptide

/product="envelop protein 1"

/notes="ORF3"

1491..2768

mat_peptide

/product="envelop protein 2"

/notes="ORF4"

2769..3419

mat_peptide

/product="non-structural protein 2"

/notes="ORF5"

3420..5312

mat_peptide

/product="non-structural protein 3"

/notes="ORF6"

5313..5474

mat_peptide

/product="non-structural protein 4a"

/notes="ORF7"

5475..6257

mat_peptide

/product="non-structural protein 4b"

/notes="ORF8"

6258..7601

mat_peptide

/product="non-structural protein 5a"

/notes="ORF9"

7602..9374

VDLCSVFLVQLFTFSPRRHWTQDNCISYPGHITCHRMANDMMNWSPTTALVV
AQLLRVPAQILMDIAGAHWGVLAIYFSWGNWAKVLVLLVLLFAGVLDANTYVTKGAA
AKGASTFAGLFPAGSRQNTIQTNGSHWINTALNCASLETGLWGLYFVTKFNSS
GCTERMAACRPLADQCGWPIYANGSGPEHRPYCWHYPKPGCGIYPAQNVGPGYVC
FTPSVNVGTTDRSCAPYNGSNDTDFVLNTRPPLGNWFGCTWMSNGFTKVCGA
PPCNIIGVGNWTLHCTPCFRKHPATYSRCSGSPNIIPRCLVDYPRYLWHVPECTLN
TLFKVMYVGGVEHRLVACNWRGRCDDDRSELSPLLSTTQWQVLVPCSFSTTL
PALTGLHQLHONIVDVQLYGVGSSI VSWAIKWEYVILLFLLADARI CSQLMMMLL
ISQAALLENLLVLAASLGRGLASLFFCFWYLRKGVPGWAGAYALYGVMPVLAH
LILALPQAYALDTMAASCGVLVGLMALTSVPYKRYISWCFWMLQYFUTRVEAH
LHWPFPLVGRGERDAVILMCMVHPTLVFDITKLLAI FGLWILQTSLLKVPYFVR
VQGLRVCVSPKMGALRYVOMAMIKVGLTCTVYVNLTPLRDWAHGLRLDLAVAE
PVFSGMETKLITWGDITACGDIINGLIPVSARRREILLGPADGNVSGKRWLLAIPIT
AYATQGTGLGCIITSLTGRDNQVEGEVQIVSTAQTFLATCINGCWTVVHAGRG
TIASPKGPVIQMYTNVDOLVGWPAPOGARSITPCTCGSSDLYLVTRHADVLPVRRRG
DRGSLSPRIPIYSLKSGSGPLHCPAGHAGVIFRAAVCTRGVAKAVPVENLETT
MESPVTDNSAPVPOFVAHLHAPTSGSKSTKVPAAYAAQGVKVLVNLPSVAATL
GFGAYMSKHGIDPNIRGTPTITTSPTITYSTYKFLADGCGSGGAYDI ILCDECHS
DATSILGIGTVLDQETAGARLVLATATPPGSVTVPHSNIEEVALSTIGEIFPYGK
ALPLEVIGGRHLIFCHSKKCDLDAALVALGINAVAYIRGLDVSIVPTGGDVVVA
TDALMTGTFDFSDVDCNTCVTQTFDFSLDPTFTIETTLTPQDAVSRQRTGRG
KPGIYRFVAPGPERPSGMDFAVLCYDAGCAWELTPTAETTVRYMNTVPLGVCQ
DHLFEWGVFTGLTHIDAFHLSQTKSGENLPYLVAOATVCARAOPPSWDQWKC
LRLKPTLHGSTPLLYRLGAVQNEVTLHPVTXYIMTMSADLEIVTSVNLVGGVLA
ALAAICSTGCVIIVGRIVSGRPILIPDOEVLYREFDEMESCQHLPIEPGMMLAE
OPKQALALLQTAHQAEVIAPOVTNNWLEAFWAKHWNFIISGQIYLAGLSTLPGN
PAIASLAPTAATVSPITTSQTLNFIILGWAAQAAPGAATAFYAGLAGRAIGSV
GLGKLVLDIAGYAGAGVALFAFKIMSGEVSTEDMNLPAILSPALVGVGVAT
ILRHVVGEGAVOMNRLIAFASRGNHVSPTHYVPEDAARVTAIILSLVTQLLR
RLHQVSESTPTPCSGWLRIIDWICELVDFKTLWKAKLMPQLPGIPLVSCQGYR
GWQDGVNHTRCHGAEITGHVKGNTMRIVGPKTCRNWSGTFPINAITYGCTPLP
APNYTALWVSAEYVEIRVDFHYVTGMTDNLRCPCQVPSPEFFTELDGVRLLR
FAPPKPLILDEVSFRVGLHDYFVGSQLPCEPEPDVAVVTSMITDPSHTAEAGRRL
ARGSPSVASSASOLSAPLKATCTTNHSDPAELIEANLLWRQMGNGNITRESEN
KVVLDSFDPLVAEDEREVSYPAEILTKRSRFAEALATWARPDPNPPILLETWKQDY
EPVHVGCDLPPQPPVPSPKKTVTLTESTVSTALAEATKSGSSSTSGIVGDS
TTTTSEPALVPSDSDAESCMPPLEGPDPLSDGWSVTSVCSQKQKVTFORQVL
SYSWTGALVTPCAEEQKLPINALSNLRLHNLVYSTTSRACQKQKVTFORQVL
DNYHQDLVKEVAAASKVKANLLSVEEACSLTPPHSARSKFGYKADVCHARKAVNH
INSWKDLLEDNVTPIITIMAKNEVFCVQPEKGRKPARLIVFDLGVVRVCEKMAIY
DIVSKLPRAVNGSSYGFQSPGQVEFLVQWKS KPTPMGFSYDTRCFDSTVTESDIR
TEALYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRCARIVLTTSCGT
LTCYKAACAAAGLRDCTMLVGGDLVVICESQVQSDAASLRAFTAMTRYVPP
RDPPEYDLELITCSNVSVAHDGAGRYVLTTRDPTPLARAWEATHTPTINSW
LGNITMFAPTLWARMILMTHFFSLMARDQLEADCEIYGCYCEPDLPLPIIQRLL
HGLSAFSLHSYSPGSEINRVAACLKGLVPLAWRHRARSVRAKLLSRGGRAAICGY
LENWAVRTKLKLTPIAAAGQLDLSGWFTAGYSGGDIYHSVSRARPRWFECULLAAG
VCIIYLLPNR"

/product="non-structural protein 5b"
/note="ORF10"

ORIGIN		Query Match	90.0%; Score 1854.8; DB 14; Length 9610;
		Best Local Similarity	93.8%; Pred. No. 0;
		Matches 1931; Conservative	0; Mismatches 127; Indels 0; Gaps 0;
QY	2	TGCGCGCTATACCGGCTATGCCAGACAGAAAGGGGCTTTTGGGATGCATAATCACCA	61
DB	3418	TGCGGCGCATACCGGCGTACGCCAGACAGAAAGGGGCTTTTGGGATGCATAATACCA	3477
QY	62	GCTTGACCGGCGGGACAAAACACAGGTGAGGGTGAGGTTTCAGATCGTCAACTGCTG	121
DB	3478	GCCTGACCGGCGGGACAAAACACAGGTGAGGGTGAGGTTTCAGATTTGTCAACTGCTG	3537
QY	122	CCGAGACTTCTTGGGCAACTGCATTAACGGGGTGTTGGACTGTCTACCATGAGGCGG	181
DB	3538	CCGAGACCTTCTTGGCAACTGCATCAACGGGGTGTTGCTGGAATGTCTACCAACGGGGCG	3597
QY	182	GAACAAGGACCAATTGCGTCACCTTAAGGGTCTTGTATCCAGATGTACCAATGTGGACC	241
DB	3598	GAACAAGGACCAATGCGTCACCTTAAGGGTCTTGTATCCAGATGTACCAATGTAGACC	3657
QY	242	AAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGGC	301
DB	3658	AAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCTGCACTTGGC	3717
QY	302	GCTCCTCGGACCTTTACCTGTCACGAGGACGCCGATGTCACTCTGTGGCGGACGGG	361
DB	3718	GCTCCTCGGACCTTTACCTGTCACGAGGACGCCGATGTCACTCTGTGGCGGCGGGG	3777
QY	362	GTGATGGGAGGGGACGCTGCTTTGCGCCCGGCTATCTTCTTACTTGAAAGGCTCCTCGG	421
DB	3778	GTGATGGGAGGGGACGCTGCTTTGCGCCCGGCTATCTTCTTACCTGAAAGGCTCCTCGG	3837
QY	422	GAGGCGCTCTGTGTGCCCGGACGACATGCCGTAGGCATATTCAGAGCGCGGTATGCA	481
DB	3838	GAGGCGCTCTGTGTGCCCGGACGACATGCCGTAGGCATATTCAGAGCGCGGTATGTA	3897
QY	482	CCCGTGGAGTGGCTAAGGCGGTGGACTTCATCCCGTAGAGAGCTTAGAGACAACCATGA	541
DB	3898	CCCGAGGATGGCTAAGGCGGTGGAAATTTGTCTCTGTAGAGAACCTTAGAGACAACCATGA	3957
QY	542	GGTCCCGCGTGTCTCAGACAACTCTCCCGACAGCAGTGCACAGAGCTACCAAGTGG	601
DB	3958	GGTCCCGCGTGTCTCAGACAACTCTGCCCGACAGCAGTGCACAGAGCTTCAGAGTGG	4017
QY	602	CCCACTGATGCTCCACCGGACGGTAAAGACCAAGGTCCCGCGCCGATACGCAG	661
DB	4018	CCCACTGATGCTCCACCGGACGGTAAAGACCAAGGTCCCGCGCTTACGGCG	4077
QY	662	CTCAGGCTACAGGTGCTGGTCTCAACCCCTCGTTGCTGCTCAACAATGGGCTTTGGTG	721
DB	4078	CTCAGGCTTACAGGTGCTGGTCTCAACCCCTCGTTGCTGCTCAACAATGGGCTTTGGTG	4137
QY	722	CTTATATGTCGAAGGCCATGGATGTATCTTAAATCAGAGCTGGGGTGAGGACAAATTA	781
DB	4138	CTTATATGTCGAAGGCCATGGATGTATCTTAAATCAGAGCTGGGGTGAGGACAAATTA	4197
QY	782	CTATGCGACCGCGATCAGTATTCACCTACGGCAAGTTCCTTCCCGACGGCGGTGTT	841
DB	4198	CCACCGGACCGCCATCAGTACTCCACTTACGGCAAGTTCCTTCCCGACGGCGGTGCT	4257
QY	842	CAGGGGTCTTTATGACATAAATTTGTGAGAGTGGCACTCCACGGATGCAATCCA	901
DB	4258	CAGGGGTCTTTATGACATAAATTTGTGAGAGTGGCACTCCACGGATGCAATCCA	4317
QY	902	TCTTGGGCAATGGCACTGCTTGTGACAAAGACAGACCGCGGGGCGAGACTGACTGTC	961
DB	4318	TCTTGGGCAATGGCACTGCTTGTGACAAAGACAGACCGCGGGGCGAGGCTGGTTGTC	4377
QY	962	TCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAACATCGAGGAGGTTG	1021

DB	4378	TCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATTTCTAATCTGAAGAGTTG	4437
QY	1022	CTCTGTCCACTACCGGAGGATGCCCTTTTATGGCAAGGCTATTCCTTCTGAAGCAATTA	1081
DB	4438	CTCTGTCCACCAACCGGGGAGATTCTCTTTTACGGCAAGGCTATCCCCCTCGAGGTATCA	4497
QY	1082	AGCGGGGAGAGCATCTCATCTTCTGCCACTCAAGAAGAGTGCGACAGGCTCGCGCAA	1141
DB	4498	AGCGGGGAGAGCATCTCATCTTCTGCCACTCAAGAAGAGTGCGACAGGCTCGCGCAA	4557
QY	1142	AACTGGTCGGTGGGCGTCAATGCGGTGCTTACTACCGCGCCTTGATGTGTCGTCA	1201
DB	4558	AACTGGTCGGTGGGCGTCAATGCGGTGCTTACTACCGCGCCTTGAGCGTGTCCGTCA	4617
QY	1202	TCCGACCAAGTGGTCAGCTGTGTGCGTGGCACTGACGCCCTCATGACGGCTTTACCG	1261
DB	4618	TCCGACCAAGTGGGACGTTGTGTGCGTGGCAACAGATGCTCTCATGACCGGCTTTACCG	4677
QY	1262	GCGACTTCGATTCGGTGTAGTACTGCAACACGTTGTTCACCCAGAGTCTGCTCCGCTACGC	1321
DB	4678	GCGACTTCGATTCGGTGTAGTATGCAACACATGTGTCAACCAGACGTCGATTTACGCC	4737
QY	1322	TTGACCTTCACTTCACTTACCATTTGAGACAATCAAGCTTCCCGAGGATGCTGTCTCCGCTACTC	1381
DB	4738	TCGACCTTACTTTTCACTTACCATTTGAGACAACACAGCTTCCCGAGGATGCTGTCTCCCGCACCC	4797
QY	1382	RACGTGGGGTAGGACTGGCAGAGGAGCCAGGATCTACAGATTTGTGGCACCGGGG	1441
DB	4798	AACGTGGGGCAGGACTGGCAGGGGAGCCAGGATCTACAGATTTGTGGCACCGGGG	4857
QY	1442	AGCGTCTTCTGGGATGTTTGAATCTGTCTCTGTGAGTGTATGACGCGGCTGTG	1501
DB	4858	AGCGGCTCTCCGATGTTTGAATCTGTGAGTGTATGACGCGGCTGTGAGCGGGCTGTG	4917
QY	1502	CTTGTATGAGCTTACGCGCGGACACACAGATTAGCTACGACATACATGAACACCC	1561
DB	4918	CTTGTATGAGCTTACACCGCGGAGACACAGATTAGCTACGACATACATGAACACCC	4977
QY	1562	GGGAGTTCCTGGTGGCCAGACCATCTTGAATTTTGGAGGGGCTCTTACGGGCTTCA	1621
DB	4978	CGGAGTTCCTGGTGGCCAGACCATCTTGAATTTTGGAGGGGCTCTTACGGGCTTCA	5037
QY	1622	CCCATAGAGCGCCACTTCTCTATCCAGACAAAGCAGAGTGGGAAAACTTCCCTATC	1681
DB	5038	CCCATATAGCGCTCACTTCTCTATCCAGACAAAGCAGAGTGGGAAAACTTCTTACC	5097
QY	1682	TGGTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGTGGGACC	1741
DB	5098	TGGTAGCGTACCAAGCCACCGTGTGCGCTAGAGCCCAAGCCCTCCCGCATCGTGGGACC	5157
QY	1742	AGATGTGAAGTGTGATCCGTCTCAAGCCACCTCCATGGGCGCAACCTCTGCTAT	1801
DB	5158	AGATGTGAAGTGTGATCCGTCTCAAGCCACCTCTTCAAGGCGCAACCTCTGCTAT	5217
QY	1802	ATAGACTGGGCGCTGCCAGAAATGAAGTCACTCTGACGACCCAGTCAACAGTATATCA	1861
DB	5218	ACAGACTGGGCGCTGTTCAAGATGAGTCACTTAAAGCAACAGTCAACAAATATCA	5277
QY	1862	TGACATGATGTGCGCTGACCTGGAGGTGCTGACGAGTACCTGGGTGCTGTTGGCGG	1921
DB	5278	TGACATGATGTGCGCTGATCTGGAGATCGTCAAGAGTACCTGGGTGCTGTTGGCGG	5337
QY	1922	TTCTGGCTGCTTGGCGCGGATTTGCCCTATCCAGGCTGGTGTGCTATAGTAGGAGA	1981
DB	5338	TTCTGGCTGCTTGGCGCGGATTTGCCCTATCCAGGCTGGTGTGCTATAGTAGGAGA	5397
QY	1982	TTGCTTCTCGGAAAGCGCGCAATCATACCCGACAGGAAAGTCTCTTACCGGAGTTCG	2041
DB	5398	TTGCTTCTCGGAGAGCGCGCAATCATACCCGACAGGAAAGTCTCTTACCGGAGTTCG	5457
QY	2042	ATGAATGGAAGTGTCT 2059	

Db	5458	ATGAGATGGAAGAGTGCT	5475	
RESULT 5				
AY615798				
LOCUS	AY615798	8791 bp	RNA	linear
DEFINITION	Hepatitis C virus isolate HCV.TWB.1 polyprotein gene, partial cds.			
ACCESSION	AY615798			
VERSION	AY615798.1	GI:48479029		
KEYWORDS	Hepatitis C virus			
SOURCE	Hepatitis C virus			
ORGANISM	Hepatitis C virus			
REFERENCE	1. (bases 1 to 8791)			
AUTHORS	Brann, T.W., Kottlilil, S., Polis, M. and Imamichi, T.			
TITLE	Identification of mutations associated with interferon resistance in HCV and HIV co-infected patients			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 8791)			
AUTHORS	Brann, T.W., Kottlilil, S., Polis, M. and Imamichi, T.			
TITLE	Direct Submission			
JOURNAL	Submitted (03-MAY-2004) LHR/CSP, SAIC-Frederick, Inc, Building 550, Room 126, 1050 Boyles Street, Frederick, MD 21702, USA			
FEATURES	Location/Qualifiers			
source	1..8791			
	/organism="Hepatitis C virus"			
	/mol_type="genomic RNA"			
	/strain="HCV1a"			
	/isolate="HCV.TWB.1"			
	/db_xref="taxon:11103"			
	/country="USA"			
	67..-8791			
	/codon_start=1			
	/product="polyprotein"			
	/protein_id="A0144336.1"			
	/db_xref="GI:48479030"			
CDS	1 translation="MSTNPKPQKTRKTRNRRPDQVFKPGGQIIVGGVYLPRRGRRL GVRATKTSRSPQRGRQIPKARPEGRWAQPGYWPWLYGNECGWAGWLLSPRG SRPQWPTPRRSLKVIDTLCGADLMGYIPLVAGPLGGAARALAHGVRVLED GUNYATNPGGCSFIPLALISCLTPASAVQVRNSGLVHTVNDPCNSIIVETAD AILHSGCICPVREGNASKWAVATVATRDGRLPSTQLRHIDLLVGSATLCSALY VGLCGSVLQGLTFSPRRHTTQDCNSIYPGHITGRHAWDMWMSPTALV AQLLRIPQAILMAGAHGVLGAIYFMSVGNWAKVLVLLFAGVDAETSTGAA ARTAFYTFAPGQNVQIITYNGSWHINRTALNCNASLDTGWLALFYNSFNS CGPERMASCRPLASPDQGPISYANGSGPEHRPYCWPYKPGIIVPAQSGVPYC FTSPVVGTTDFKGYPTYNWANDTDVFLANTRPPLGNWFGCTWNSGTGFKVCGA PPCVIGVGNTHLCTDCPKRHPATYSRCSGPMITPRCLVDYFYRLWHYPTCLNY TLFKRMVYGVGVEHREAAACNWTGRERCDLDDRSLSPLLLITTTQWQVLPSCFTL PALTTGLIHLQNIIVQYLVGVSSIVSWAIKWEYVILLFLLADARICSLMMWL ISOAEALLENLIVLNAASLAGTHLASFLVFFCFAWYLGKWPVGNMAYAFPGWPLLL LLIALPQRAVALDTEMAACGGVVLVGLMALTLSPHYKRYISMCLWMLQVFLTRAEAL LQWVPLVGRGDAVILLVCVHPALVVDITKLIALALGPLWILQTSLLKVPYVR VQGLIRICALRAGGHVQMAIIKAGALTGYIYNHITPLDMHAGNLRDLAVAVE PVFVSQMETKLTITWADTAACGDIINGLPVSARRREIILGPADGWSKGRLLAPIT AYAAQRTGLGCIITLGRDNQVEGEIIVSTAAQTFLATCINGCVWTVYHAGTR TIASPKGPVOMYTNIDQLVGPAPQAGHARSLPCTCGSSDLVLTVRHADVIPRRRG DSRGSLSPRIISYLGSSGGLLCPAGHARSLPCTCGSSDLVLTVRHADVIPRRRG MRSPVTDNSPPAVPQSFQVAHLHAPTSGSKTKVPAAYAGQYKVLNPSVAATL GPGAYMSKAGHDNPTVGTITITGSPITITGRLADGGCGGAYDIIICDECHS TDATSLIGTVDLQDRTAGRLVLAATTPGSIIVPHNIEVALSTSLTFIPFGK AIPLEAIKGRHLIFCHSKKCDLAKLVALGINAVAYRGLDVSIVPTSGDVVYA TDALMTGFTGDPSVIDNCVTQVDFSLDPTFTIETTLPODASVTRQRRGTGRG KPGIYRFAGERPSGMFSSVLCVCDAGCANVELTPAETTVLRARVMTPLGVQ DHELFWEVFTGLTIDAHPLSQTKSGENLPVLYATVCAQAPPPSWDMWKC LRLKPTLHPTPLRLVAGVQNEVITLTPVTKIIMTMSADLEVSTISLVGGVLA ALQAYCLSTGCWVIGRIVLSGPAVIPDEVLRYFDENECSQHLPIYEQMGLAE OFKQALGLIQASRAQVIAVAPQVNWQLEAFWAKHWNFIISGQIYLAGLSTLGN PAIASLMAVTAAPVPLTSTQTLNFIILGSWAAQLAAGAAAFVAGLAGLAAGIAGV GLCKVLVDIILAGYAGVAGLVAFKIMSGEPPTEDLVNLPAILSPGALVGVCAV ILRHVGPGEVAGVQWNRLLIAPASRGNHVSPTHYPESDAARVTAIISLTVXLLR RLHWVSSECTTTPCDSWLRDITWDICVETLWFKLAKMLPQLPGIPVSCQRGR GVRWGDGIMHTRCHGABITGVHKNGTMRIVGPKTCRNWMSGTFPFINAYTGPCTPLP			

APNTYFALWRVSAEYVEIRVGDPHYVTGMDNLKCPQVSPSEFFTELDGVRLLR				
FAPCPKLLRDEVSFRVGLHDYFVGSQPLPCPEPDAVLITSLMTDPSHITAGARRRL				
GAQSPSSASSASQLSAPSRATCTTNHSDPSDAELIENLIRWEMGNITRVSEEN				
KVVIDLSPDLVAEEDREVSAAEILKRSRFTPALPWAERLPNPLLESWKTPDY				
EPPTVHGCPPLPPQSPVPVPPRKRKTIVLSBSTVSTALAEATKSGFSSTGTTGDN				
TASSSEPSPICLPDSDAESCMPPLEGDPDLSDGSWSTVSEACQTKVTFDRLQVL				
SYSWGALVTPCAAEOKLPINALSNLSLLRHNLVYVSTTSACOROKQKVDVRCHAKVNH				
DSHYQDVILKEVKAASKVANLLSVEACSLTPSPHAKSKFGYKADVRCHARKVNH				
INSWKDILLEASVTPIDITIMAKNEVFQVPEKGGCKPARLIVFDLGVVRVFEKMDLY				
DVSKLPLAVMGSYSYGFQYQRFVFLQAWKSKTTPMGFSYDFRFDSTVTFESDIR				
LTCTYKAAQACRAAGRLDCTWXXGDDLVVICESQVOBDAASLRAFTAMTRYSSAPP				
GDPPPEYDLELITSCSNVSVAHGAGKRVVYLRDPTTPLARAWEATARTHPVNSW				
LGNIIMFAPLTHSVMILMTHFSPVLIARDQLEQALDCEIYGACYSIEPLDLPILORL				
HGLSAFSLHSYSPGGINRVAAC"				
ORIGIN				
Query Match	89.6%	Score 1846.4;	DB 14;	Length 8791;
Best Local Similarity	93.5%	Pred. No. 0;		
Matches 1925;	Conservative 1;	Mismatches 132;	Indels 0;	Gaps 0;
QY	2	TGGCGCCTATCACGGCCTATGCCAGCAGACAAGGGCGCTTTTGGGATCATATAATCACCA	61	
DB	3143	TGGCGCCCATCACGGCGTATGCCARCAAGGGCGCTCTTAGGATCATATAATCACCA	3202	
QY	62	GCTTGACCGCGCGGACAAAAACACAGGTGGAGGGTGAGTTTCTAGATCGTCACTGCTG	121	
DB	3203	GTCTGACCGCGCGGACAAAAACACAGGTGGAGGGTGAGATTGTGTCACTGCTG	3262	
QY	122	CCGAGACTTTCTTGGCAACTGCATTAACGGGCTGTGTGGACTGTCTACCATGGAGCGG	181	
DB	3263	CCGAGACTTTCTTGGCAACTGCATTAACGGGCTGTGTGGACTGTCTACCATGGAGCGG	3322	
QY	182	GAACAAGGACCACTTGCGTCACCTAAGGGCTCTGTATCCAGATGTACACCAATGTGGACC	241	
DB	3323	GAACAAGGACCACTTGCGTCACCTAAGGGCTCTGTATCCAGATGTATATAATAGACC	3382	
QY	242	AGAAGCTCTAGGCTGGCGCGCTCCCAAGGGTCCCGCTCATTAACCATGACCTTGG	301	
DB	3383	AGAAGCTCTAGGCTGGCGCGCTCCCAAGGGTCCCGCTCATTAACCATGACCTTGG	3442	
QY	302	GCTCTCTCGGACCTTTACCTGTGTACGAGGACGCGCGATGTCTCTGTGCGCGGACGGG	361	
DB	3443	GCTCTCTCGGACCTTTACCTGTGTACGAGGACGCGCGATGTCTCTGTGCGCGGACGGG	3502	
QY	362	GTGATGGCAGGGGACGCTGCTTTGCGCGCGCTATCTCTTACTTTGAAAGGCTCTCGG	421	
DB	3503	GTGATGGCAGGGGACGCTGCTTTGCGCGCGCTATCTCTTACTTTGAAAGGCTCTCGG	3562	
QY	422	GAGGCGCTCTGTGTGCGCGCGGACGATGCGGTAGGCATATTCAGAGCGCGCGTATGCA	481	
DB	3563	GAGGCGCTCTGTGTGCGCGCGGACGATGCGGTAGGCATATTCAGAGCGCGCGTATGCA	3622	
QY	482	CCGCTGGAGTGGCTAAGCGGCTGGACTTCTCCCGGTAGAGAGCTTAGAGACAACCATGA	541	
DB	3623	CCGCTGGAGTGGCTAAGCGGCTGGACTTCTCCCGGTAGAGAGCTTAGAGACAACCATGA	3682	
QY	542	GGTCCCGGCTGTCTCAGACAACCTCTCCCGACACAGCAGTGGCCAGAGCTACCAAGTGG	601	
DB	3683	GGTCCCGGCTGTCTCAGACAACCTCTCTCCCGACCGGGGTGGCCCGAGAGCTTCCAGTGG	3742	
QY	602	CCCACTGCTGCTCCCAACCGGCGGTAAGAGCCAGAGGTCCCGCGCGCATACGCAG	661	
DB	3743	CCCACTGCTGCTCCCAACCGGCGGTAAGAGCCAGAGGTCCCGCGCGCATACGCAG	3802	
QY	662	CTCAGGGCTACAGGCTGGTGTCTCAACCCCTCCGTCTGTCTGCAACAATGGGCTTTGGTG	721	
DB	3803	CTCAGGGCTACAGGCTGGTGTCTCAACCCCTCCGTCTGTCTGCAACAATGGGCTTTGGTG	3862	
QY	722	CTTACATGTCCAAAGGCCCATGGATTGATCTTAACATCAGGACTGGGGTGAGGACAAATTA	781	
DB	3863	CTTATATGTCCAAAGGCCCATGGATTGATCTTAACATCAGGACTGGGGTGAGGACAAATTA	3922	

```

782  CTACTGCGCCCGGATCACGATATTCCACTACCGCAAGTTCTTTGCGCGCGGGTGT 841
    |||||
3923  CTACTGCGCCCGGATCACGATATTCCACTACCGCAAGTTCTTTGCGCGCGGGTGT 3982
    |||||
842  CAGGGGTGTTATGACATAATAATTGTGACAGTGCCACTCCACGGATGCAACATCCA 901
    |||||
3983  CAGGGGTGTTATGACATAATAATTGTGACAGTGCCACTCCACGGATGCAACATCCA 4042
    |||||
902  TCTTGGGCATTGGCATCTGCTTACCAAGCAGAGACCGGGGGGCGAGACTGACTGTGC 961
    |||||
4043  TCTTGGGCATTGGCATCTGCTTACCAAGCAGAGACTGCGGGGCTAGACTGGTTGTGC 4102
    |||||
962  TCGCCACCGCTACCCCTCCGGGCTCCGTCACGTGTGCCCATCTCAACATCGAGGAGTTG 1021
    |||||
4103  TCGCCACCGCTACCCCTCCGGGCTCCATCACTGTGCCCATCTCAACATCGAGGAGTTG 4162
    |||||
1022  CTCGTGCTCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAAGCAATTA 1081
    |||||
4163  CTCGTGCTCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAGGCAATCA 4222
    |||||
1082  AGGGGGGAGACATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1141
    |||||
4223  AGGGGGGAGACATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4282
    |||||
1142  AACTGTCTCGGTTGGGCGTCAATGCGGCTTACTACCGCGGCTTGTATGTGCTGCTCA 1201
    |||||
4283  AACTGTCTCGGTTGGGCGTCAATGCGGCTTACTACCGCGGCTTGTATGTGCTGCTCA 4342
    |||||
1202  TCCCGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1261
    |||||
4343  TCCCGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4402
    |||||
1262  GCGACTTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1321
    |||||
4403  GCGACTTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4462
    |||||
1322  TTGACCTTACCTTCACATTTGAGACAAATCACGCTTCCCGAGATGCTGCTCCCGTACTC 1381
    |||||
4463  TTGACCTTACCTTCACATTTGAGACAAATCACGCTTCCCGAGATGCTGCTCCCGTACTC 4522
    |||||
1382  AACTGCGGGTGTAGCTGCGCAGAGGAAACCGAGGATCTACAGATTTGTGGACCGGGGG 1441
    |||||
4523  AACTGCGGGTGTAGCTGCGCAGAGGAAACCGAGGATCTACAGATTTGTGGACCGGGGG 4582
    |||||
1442  AGCGCTCTCTGCGCATGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
    |||||
4583  AGCGCTCTCTGCGCATGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4642
    |||||
1502  CTTGCTATGAGCTTACGCGCGCGGAGACCAAGCTAGGCTACGAGCATACATGAACACC 1561
    |||||
4643  CTTGCTATGAGCTTACGCGCGCGGAGACCAAGCTAGGCTACGAGCATACATGAACACC 4702
    |||||
1562  CCGGACTTCCCGTGTGCGCAAGACCAATTTTGAATTTTGGAGGGCGTCTTTACGGTCTCA 1621
    |||||
4703  CCGGCTTCCCGTGTGCGCAAGACCAATTTTGAATTTTGGAGGGCGTCTTTACGGTCTCA 4762
    |||||
1622  CCCACATAGACCGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
    |||||
4763  CCCACATAGACCGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCTTATC 4822
    |||||
1682  TGGTAGCGTACCAAGCAGCGGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGCTGGGACC 1741
    |||||
4823  TGGTAGCGTACCAAGCAGCGGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGCTGGGACC 4882
    |||||
1742  AGATGTGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 1801
    |||||
4883  AGATGTGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 4942
    |||||
1802  ATAGACTGGGCGCTGTGCAAGATGAGTCAACCTGCGCAGCAGCCAGTCAACAGTATATCA 1861
    |||||
4943  ACAGGCTGGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTATCA 5002
    |||||
1862  TGACATGTATGTGCGGCTGACCTGGAGGTCGTACAGAGTACCTGGGTCGTGCTGCGGGG 1921
    |||||
```

```

Db      5003  TGACATGTATGTGCGGCGACCTGGAGGTCGTACAGAGTACCTGGGTCGTGCGGTCGGC 5062
    |||||
QY      1922  TTCTGGCTGCTTTTGGCCGCGTATTCCCTATCCACAGGTCGCTGCTGCTATAGTAGGTA 1981
    |||||
Db      5063  TTCTGGCTGCTTTTGGCCGCGTACTCCCTATCCACAGGTCGCTGCTGCTATATAGGACGA 5122
    |||||
QY      1982  TTGTCTTGTTCGGAAAGCGGCAATCATACCGCAGGGAAGTCTCTTACCGGGAGTTCCG 2041
    |||||
Db      5123  TTGTCTTGTTCGGGAAGCGGCGAGTCATACCGCAGGAGGTCCTCTACCGAGAGTTCCG 5182
    |||||
QY      2042  ATGAAATGGAAGAGTGCT 2059
    |||||
Db      5183  ATGAAATGGAAGAGTGCT 5200
    |||||

RESULT 6
AXI64584
LOCUS   AXI64584               6299 bp      DNA      linear      PAT 22-JUN-2001
DEFINITION   Sequence 6 from Patent WO0138360.
ACCESSION   AXI64584
VERSION     AXI64584.1   GI:14545518
KEYWORDS    .
SOURCE      synthetic construct
ORGANISM    synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Coit,D.C., Medina-Selby,A.C., Selby,M.C. and Houghton,M.C.
  TITLE     Novel hcv non-structural polypeptide
  JOURNAL   Patent: WO 0138360-A 6 31-MAY-2001;
            CHIRON CORPORATION (US)
FEATURES    Location/Qualifiers
             1..6299
             /organism="synthetic construct"
             /mol_type="unassigned DNA"
             /db_xref="taxon:32630"
             /note="pNS34a"
             1990..4050
             /note="unnamed protein product"
             /codon_start=1
             /transl_table=11
             /protein_id="CAC42724.1"
             /db_xref="GI:14545519"
             /translation="MADTAYAOOTRGLGCIITSLTGRDKNOVEGEVOIVSTRAQTF
LATCINGVCVTYHAGATRTIASPKVPIQMTNVDQLVGPASQGRISRTPTCTGS
SDLXIVTRHADVIPIVRRGDSRSLSPRIISYILKSGGGPLCPAGHAVGIFRAVC
TRGKAVADFIPIVENLETTMSPVTDNSPVPVQSFVAHLHAPTSKSTKVPAA
YAAQGYKVLNPSVAATLGFAYMSKAHGDIPNIRTVRTITTSPTTYTGKFLA
DGGCGGAYDIIICDECHSDATSIIGITVLDQAEATAGARLVLATATPGSVTPH
PNIEVALSTTGEIPIFYKAIPLFVIGGRHLIFCHSKKCDLAALVALGINAVAY
YRGDVSVIPFGSDVVVVVATDALMTGTGDEFSDVIDCNCVTQIVDFSLDPTFTETI
TLPQDASVTRORRTGRKPGFIYRFVAPGRPSGFMFDSVLCBCYDAGCAWELTPA
ETTVLRAYMNTPLGVQDHLFEWGVFTGLTHIDAHFLSQSGENLPLVAYQA
TVCAQAAPPSPQDMWKLIIRLKPGLPTLLYRLCAVQNEITLTPVTKYIMTCM
SADLEVTWTWLVGGVLAALAAVCLSGCVVIVGRVVLGKPAIIPREVLYREFDE
MEEC"
CDS
             1..6299
             /organism="synthetic construct"
             /mol_type="unassigned DNA"
             /db_xref="taxon:32630"
             /note="pNS34a"
             1990..4050
             /note="unnamed protein product"
             /codon_start=1
             /transl_table=11
             /protein_id="CAC42724.1"
             /db_xref="GI:14545519"
             /translation="MADTAYAOOTRGLGCIITSLTGRDKNOVEGEVOIVSTRAQTF
LATCINGVCVTYHAGATRTIASPKVPIQMTNVDQLVGPASQGRISRTPTCTGS
SDLXIVTRHADVIPIVRRGDSRSLSPRIISYILKSGGGPLCPAGHAVGIFRAVC
TRGKAVADFIPIVENLETTMSPVTDNSPVPVQSFVAHLHAPTSKSTKVPAA
YAAQGYKVLNPSVAATLGFAYMSKAHGDIPNIRTVRTITTSPTTYTGKFLA
DGGCGGAYDIIICDECHSDATSIIGITVLDQAEATAGARLVLATATPGSVTPH
PNIEVALSTTGEIPIFYKAIPLFVIGGRHLIFCHSKKCDLAALVALGINAVAY
YRGDVSVIPFGSDVVVVVATDALMTGTGDEFSDVIDCNCVTQIVDFSLDPTFTETI
TLPQDASVTRORRTGRKPGFIYRFVAPGRPSGFMFDSVLCBCYDAGCAWELTPA
ETTVLRAYMNTPLGVQDHLFEWGVFTGLTHIDAHFLSQSGENLPLVAYQA
TVCAQAAPPSPQDMWKLIIRLKPGLPTLLYRLCAVQNEITLTPVTKYIMTCM
SADLEVTWTWLVGGVLAALAAVCLSGCVVIVGRVVLGKPAIIPREVLYREFDE
MEEC"
ORIGIN
Query Match      86.7%; Score 1787; DB 6; Length 6299;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY      1  ATGGCGCCTATCAGGCGCTATGCCAGCAGACAGAGGGCCCTTTGGGATGATATATCACC 60
    |||||
Db      1990  ATGGCGCCCATCAGGCGGTACGCCAGCAGACAAAGGGCCCTCCTAGGGTGATATATCACC 2049
    |||||
QY      61  AGCTTGACCGCGCGGACAAAACACAGGTGAGGCTGAGGTTTCAGATCGTGTCAACTGCT 120
    |||||
Db      2050  AGCCTTACTGCGCGGACAAAACCAAGTGAGGGTGAGGTCAGATTTGTCACTGCT 2109
    |||||
QY      121  GCCCAGACTTTCTTTGGCAACCTGATTAACCGGGGTGTGTGGACTGTCTACATGAGACC 180
    |||||
Db      2110  GCCCAACCTTCTCTGGCAACCTGATCAATGGGGTGTGCTGAGCTGTCTACACCGGGGCC 2169
    |||||
```

QY 181 GGAACAAGGACCAATTGGCTCACTAAGGGTCTCTGTATCCAGATGTACCAATATGGAC 240
DB 2170 GGAACGAGGACCAATCGGCTCAACCAAGGGTCTGTATCCAGATGTATACCAATGTAGAC 2229
QY 241 CAGACCTCTAGGTGGCGGCTCCCAAGGTGGCGCTCATTAACACCATGACTTGC 300
DB 2230 CAGACCTTGGGCTGGCGGCTTCGCAAGGTACCGCTCATTTGACCCCTGCACTTGC 2289
QY 301 GGTCTCTCGACCTTTACCTGGTCAAGGACGCGCATGCTCATCTGTGCGCCGACGG 360
DB 2290 GGCTCTCGACCTTTACCTGGTCAAGGACGCGCATGCTCATCTGTGCGCCGACGG 2349
QY 361 GGTGATGGCAGGGCAGCTGCTTTGCGCCCGGCTATCTCTTAATGAAGGCTCTCTG 420
DB 2350 GGTGATGACAGGGGAGCTGCTGCTGCGCCCGGCTATCTCTTAATGAAGGCTCTCTG 2409
QY 421 GGAGGCCCTCTGCTGCGCCGCGGAGACATGCGGTAGGCATATTCAGAGCGCGGTATGC 480
DB 2410 GGGGGTCCGCTGTTGTGCGCCGCGGAGACATGCGGTAGGCATATTCAGAGCGCGGTATGC 2469
QY 481 ACCGCTGGAGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGAGCTTTAGAGCAACCATG 540
DB 2470 ACCGCTGGAGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGAGCTTTAGAGCAACCATG 2529
QY 541 AGGTCCCGGTGTTCTCAGACAACTCTCCCAAGGAGTGGCCGAGAGCTACCAAGTG 600
DB 2530 AGGTCCCGGTGTTCTCAGACAACTCTCTCCAGGAGTGGCCGAGAGCTTCCAGGTG 2589
QY 601 GCGCACTGATGCTCCACGCGGAGCGGTAAAGACACCAAGTTCGCGCGCATAGCA 660
DB 2590 GCTCACCTCATGCTCCACAGGAGCGGCAAGACCAAGTTCGCGCGCATATGCA 2649
QY 661 GCTCAGGGCTACAGGTGCTGCTCAACCCCTCGTTCGCTCAACATAGGCTTTGGT 720
DB 2650 GCTCAGGGCTACAGGTGCTGCTCAACCCCTCGTTCGCTCAACATAGGCTTTGGT 2709
QY 721 GCTTACATGCTCAAGGCCATGGATGATCTTAACATCAGAGTGGGGTGAGGACAAAT 780
DB 2710 GCTTACATGCTCAAGGCCATGGATGATCTTAACATCAGAGTGGGGTGAGGACAAAT 2769
QY 781 ACTACTGGACCGGATCAGCTATTCACCTAGGCAAGTTCCTTGGCGAGCGGGTGT 840
DB 2770 AACACTGGACCGGATCAGCTATTCACCTAGGCAAGTTCCTTGGCGAGCGGGTGT 2829
QY 841 TCAGGGGCTCCTATGACATAAATTTGTGACGAGTGCCACTCCAGGATGCAACATCC 900
DB 2830 TCAGGGGCTCCTATGACATAAATTTGTGACGAGTGCCACTCCAGGATGCAACATCC 2889
QY 901 ATCTTGGGCAATGGCACTGCTTGTGACCAAGCAGAGACCGCGGGGCGAGACTGACTGTG 960
DB 2890 ATCTTGGGCAATGGCACTGCTTGTGACCAAGCAGAGACTGCGGGGCGAGACTGACTGTG 2949
QY 961 CTGCGCACTGCTACCTCGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 2950 CTGCGCACTGCTACCTCGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3009
QY 1021 GCTCTGTCTACCTACCGAGAGATCCCTTTTATGCGAAGGCTATTCCTTGAAGCAAT 1080
DB 3010 GCTCTGTCTACCTACCGAGAGATCCCTTTTATGCGAAGGCTATTCCTTGAAGCAAT 3069
QY 1081 AAGGGGGGAGACATCTCATCTTCTGTGATCAAGAGAGTGGCGAGCTGCGCGCA 1140
DB 3070 AAGGGGGGAGACATCTCATCTTCTGTGATCAAGAGAGTGGCGAGCTGCGCGCA 3129
QY 1141 AAATGCTGCTGGCGGTCAATGCGGTGGCTTACTACCGCGCTTGTATGCTCGCT 1200
DB 3130 AAGTGTGCTGCTGGCGGTCAATGCGGTGGCTTACTACCGCGCTTGTATGCTCGCT 3189
QY 1201 ATCCGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 3190 ATCCGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3249

QY 1261 GGGCACTTCGATTCGGTGTATGACTGCAACACGTTGTGTCAACGACAGCTCGACTTCAGC 1320
DB 3250 GGGCACTTCGATTCGGTGTATGACTGCAACACGTTGTGTCAACGACAGCTCGACTTCAGC 3309
QY 1321 CTTGACCTTACCTTCCACATTCAGACCAATCAGCTTCCCGAGAGTGTCTCTCCCGTACT 1380
DB 3310 CTTGACCTTACCTTCCACATTCAGACCAATCAGCTTCCCGAGAGTGTCTCTCCCGTACT 3369
QY 1381 CAACGCTCGGGTAGGACTGGCAGAGGGAAGCCAGGCACTCTACAGATTTTGTGGCACCAGG 1440
DB 3370 CAACGCTCGGGTAGGACTGGCAGAGGGAAGCCAGGCACTCTACAGATTTTGTGGCACCAGG 3429
QY 1441 GAGCGTCTTCTGCGCATGTTTGAATGCTGCTCTCTGCGAGTGTATGACGCGGGTGT 1500
DB 3430 GAGCGCTCTCTGCGCATGTTTGAATGCTGCTCTCTGCGAGTGTATGACGCGAGCTGT 3489
QY 1501 GCTTGTATGACTTACGCGCCGCGAGACCAAGTGTAGGCTACGAGCATACAGACAC 1560
DB 3490 GCTTGTATGACTTACGCGCCGCGAGACTACAGTGTAGGCTACGAGCATACAGACAC 3549
QY 1561 CCGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCGCTCTTTAGCGGTCTC 1620
DB 3550 CCGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCGCTCTTTACAGGCTC 3609
QY 1621 ACCCATAGACCGCCACTTCTATCCCAAGCAAAAGCAGAGTGGGGAACCTTCCCTAT 1680
DB 3610 ACTCATATAGATGCCACTTCTATCCCAAGCAAAAGCAGAGTGGGGAACCTTCCCTAT 3669
QY 1681 CTGTAGCGTACCAAGCCAGCTGCTGCGCTAGAGCTCAAGCCCTCCCGGTGGGAC 1740
DB 3670 CTGTAGCGTACCAAGCCAGCTGCTGCGCTAGAGCTCAAGCCCTCCCGGTGGGAC 3729
QY 1741 CAGATGTGAAGTGTCTGATCCGTCTCAAGCCCAACCTCTCAATGGGCAACACCTCTGCTA 1800
DB 3730 CAGATGTGAAGTGTCTGATCCGTCTCAAGCCCAACCTCTCAATGGGCAACACCTCTGCTA 3789
QY 1801 TATAGCTGGGCGCTGTCCAGATGAAGTCACTGACGCAACCTGAGTCAACAGTATATC 1860
DB 3790 TACAGCTGGGCGCTGTCCAGATGAAGTCACTGACGCAACCTGAGTCAACAGTATATC 3849
QY 1861 ATGACATGTATCTGCGCTGACCTGGAGTGTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTG 1920
DB 3850 ATGACATGTATCTGCGCTGACCTGGAGTGTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTG 3909
QY 1921 GTTCTGGCTGCTTGGCGCGTATGCTTATCCATCAAGGCTGCTGCTGCTGCTGCTGCTGCTG 1980
DB 3910 GTTCTGGCTGCTTGGCGCGTATGCTTATCCATCAAGGCTGCTGCTGCTGCTGCTGCTGCTG 3969
QY 1981 ATTGCTTGTCCGGAAGCCGCAATCATACCCGACAGGGAAGTCTCTTACCGGGAGTTC 2040
DB 3970 GTCTGCTTGTCCGGAAGCCGCAATCATACCCGACAGGGAAGTCTCTTACCGGGAGTTC 4029
QY 2041 GATGAATGGAAGTGTCT 2059
DB 4030 GATGAATGGAAGTGTCT 4048

RESULT 7

106434

LOCUS

DEFINITION

Sequence 48 from Patent EP 0318216.

ACCESSION

I06434

VERSION

I06434.1

KEYWORDS

GI:590311

SOURCE

Unknown.

ORGANISM

Unkown.

REFERENCE

1 (bases 1 to 5360)

AUTHORS

Houghton, M., Choo, Q.-L. and Kuo, G.

TITLE

Nanov diagnostics and vaccines

JOURNAL

Patent: EP 0318216-A1 48 31-MAY-1989;

FEATURES

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

Unknown.

Unkown.

1 (bases 1 to 5360)

Houghton, M., Choo, Q.-L. and Kuo, G.

Nanov diagnostics and vaccines

Patent: EP 0318216-A1 48 31-MAY-1989;

Location/Qualifiers

1. .5360

source

Sequence 48 from Patent EP 0318216.

I06434

I06434

I06434.1

GI:590311

ORIGIN		/organism="unknown" /mol_type="unassigned DNA"		Query Match Best Local Similarity Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;		
QY	2	TGGCGCTATCAGCGCTATGCCAGACAGACAAGGGCGCTTTGGGATGCATATACCA	61	1891	TCGCCACCGCACCCCTCGGGCTCCGTCACCTGTGCCCATCCCAACATCGAGGAGTTG	
DB	931	TGGCGCCCATCAGCGGTAGCCAGACAGACNAGGGGCTCTTAGGGTGCAATACCA	990	QY	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTGAAGCAATTA
QY	62	GCTTGACCGCGCGGACAAAAACCAAGGTGAGGGTTCAGATCGTGCACACTGCTG	121	DB	1951	CTCTGTCCACCACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTGAAGTAAATCA
DB	991	GCCTAACTGGCGGACAAAAACCAAGTGGAGGGTGAGTCCAGATTGTGCAACTGCTG	1050	QY	1082	AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGAGTGCAGAGCTCGCGCAA
QY	122	CCAGACTTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCGG	181	DB	2011	AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGAGTGCAGAGCTCGCGCAA
DB	1051	CCCAAACTTCTTGGCAACCTGCATCAATGGGGTGTGTGGACTGTCTACCACGGGCGG	1110	QY	1142	AACTGGTCCGTTGGGGCGTCAATGCCGTGGCTTACTACCGGGCGCTTGTATGTCTCCGTCA
QY	182	GAACAGGACCATTCGCTCACTAAGGGTCTGTATCCAGATGTACCAATGTGGACC	241	DB	2071	AGCTGGTCCGATTTGGGCATCAATGCCGTGGCTACTACCGCGGTCTTGTAGCGTCCGTCA
DB	1111	GAACAGGACCATTCGCGGTCAACCAAGGGTCTGTATCCAGATGTATACCAATGTAGCC	1170	QY	1202	TCCCGACACAGTGGTGTGCTGTGCGTGGCAACTGAGCGCCCTCATGACCGGCTTTACCG
QY	242	AAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGG	301	DB	2131	TCCCGACACAGCGGCGATGTGTGCTGTGCGAACCGATGCCCTCATGACCGGCTATACCG
DB	1171	AAGACCTTGTGGCTGGCGCGCTCCGCAAGGTAGCGCTCATTTGACACCTTGCACTTGG	1230	QY	1262	GCGACTTCGATTCGGGTAGACTGCAACACGCTGTGTCAACAGAGTGTCAACCGGCTTTACCG
QY	302	GCTCTCGACCTTTACCTGGTCAAGGACCGCGATGTCAATCTGTGCGCGCAAGCGG	361	DB	2191	GCGACTTCGATTCGGGTAGACTGCAACACGCTGTGTCAACAGAGTGTCAACCGCTTCAGCC
DB	1231	GCTCTCGACCTTTACCTGGTCAAGGACCGCGATGTCAATCTGTGCGCGCAAGCGG	1290	QY	1322	TTGACCTTACCTTCAACATTGAGACAATCAACGCTTCCCGAGGATGCTGTCTCCGTACTC
QY	362	GTGATGGAGGGGACGCTTGTGCGCGCGCTTCTTACTTTGAAAGGCTCTCGG	421	DB	2251	TTGACCTTACCTTCAACATTGAGACAATCAACGCTTCCCGAGGATGCTGTCTCCCGCACTC
DB	1291	GTGATGGAGGGGACGCTTGTGCGCGCGCGCTTCTTACTTTGAAAGGCTCTCGG	1350	QY	1382	RACGTCCGGGTAGGACTGGCAGAGGAGCCAGGCACTACAGATTTGTGGCACCGGGGG
QY	422	GAGGCGCTCTGCTGTGCCCGGAGGACATGCGGTAGGCATATTACAGAGCGCGGTATGCA	481	DB	2311	AACGTCCGGGGCAGGACTGGCAGGGGAGCCAGGCACTACAGATTTGTGGCACCGGGGG
DB	1351	GGGGTCCGCTGTGTGCCCGCGGGGACGCGGTGGGCAATTTTAGGGCCCGGTGTGCA	1410	QY	1442	AGCGTCTTCTGGCATGTTTGACTCGTCTGTCTCTCGAGTGTCTATGACGCGGGTTGTG
QY	482	CCGCTGGAGTGGCTAAGCGGTGGACTTCAATCCCGGTAGAGAGCTTTAGAGACAACATGA	541	DB	2371	AGCGCCCTTCCGGCATGTTTGCATCGTCCGTCTCTGTGAGTGTCTATGACGAGGCTGTG
DB	1411	CCGCTGGAGTGGCTAAGCGGTGGACTTCAATCCCTGTGGAGAACCTAGAGACAACCATGA	1470	QY	1502	CTTGGTATGAGCTTACGCGCCCGAGACCAAGTGTAGGCTACGAGCATACATGAACACCC
QY	542	GTGTCGCGGTGTCTCAGACAACTCTCCCGACAGAGTCCCGCGCATACCGAGCTACCAAGTGG	601	DB	2431	CTTGGTATGAGCTACGCGCCCGCAGACTACAGTGTAGGCTACGAGCGTACATGAACACCC
DB	1471	GGTCCCGGTGTCTCAGGATAACTCTCTCTCACCAAGTAGTCCCCAGAGCTTCCAGGTGG	1530	QY	1562	CGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCGCTTTTACCGGGTCTCA
QY	602	CCACCTGCATGCTCCACCGGAGCGGTAAAGACACCAAGGTCCCGCGCATACGCAG	661	DB	2491	CGGGCTTCCCGTGTGCCAGACCATCTTGAATTTTGGGAGGGCGCTTTTACAGGGCTCA
DB	1531	CTCACCTCCATGCTCCACAGGACGGGCAAAAGCACCAAGGTCCCGGTGCATATGCAG	1590	QY	1622	CCCATAGAGCGCCCACTTCTCTATCCAGACAAAGCAGAGTGGGGGAAAACTTCCCTATC
QY	662	CTCAGGGCTAAGGTGTGTGCTCAACCCCTCGTTGCTGCAACATGGGGTTTGGTG	721	DB	2551	CTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC
DB	1591	CTCAGGGCTAATAAGGTGTGTGCTCAACCCCTCTGTGCTGCAACACTGGGGCTTTGGTG	1650	QY	1682	TGGTAGCGTACCAAGCCACCGTGTGCGTGTAGAGTCAAGCCCTTCCCGGTGTGGGACC
QY	722	CTTACATGTCCAGGCGCATGGATGATCCTTAACATCAGGCTGGGGTGAGCAATTA	781	DB	2611	TGGTAGCGTACCAAGCCACCGTGTGCGTGTAGGCTCAAGCCCTTCCCGCATCGTGGGACC
DB	1651	CTTACATGTCCAGGCGCATGGATGATCCTTAACATCAGGCTGGGGTGAGCAATTA	1710	QY	1742	AGATGTGGAAGTGTGTTGATTCGCTCTCAAGCCCACTCTCATATGGGCGCAACCTCTGCTAT
QY	782	CTACTGGGCGCCGATCAGTATTTCCACTACGGCAAGTTCCTTCCGACGGCGGGTGT	841	DB	2671	AGATGTGGAAGTGTGTTGATTCGCTCTCAAGCCCACTCTCATATGGGCGCAACCTCTGCTAT
DB	1711	CCACTGGGAGCCCATCAGTACTCCACTACGGCAAGTTCCTTCCGACGGCGGGTGT	1770	QY	1802	ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGCAACCAAGTCAACAAATATCA
QY	842	CAGGGGTGTATGACATAATAATTTGTGACAGTGCATCTCCACGGATGCACATCCA	901	DB	2731	ACAGACTGGGCGCTGTTTCAGAAATGAATCAACCTGACGCAACCAAGTCAACAAATATCA
DB	1771	CAGGGGTGTATGACATAATAATTTGTGACAGTGCATCTCCACGGATGCACATCCA	1830	QY	1862	TGACATGTATGTGCGGTGACCTGGAGGTGCTGACAGTACCTGGGTGTCTGTGGCGGG
QY	902	TCTTGGGCTTGGCATCTGCTTACCAAGAGAGACCGGGGGGAGACTGACTGTGC	961	DB	2791	TGACATGTATGTGCGCGACCTGGAGGTGCTGACAGCACCTGGGTGTCTGTGGCGGG
DB	1831	TCTTGGGCTTCGGCATCTGTCTTGACCAAGAGAGACTGCGGGGGGAGACTGCTGTGC	1890	QY	1922	TTCTGGCTCTTGGCGCGTATTTGCCATTCACAGGCTGCGTGGTGTATAGTAGGTAGGA
QY	962	TCGCCACCGCTACCCCTCCGGGCTCCGTCACCTGTGCCCATCTTAACATCGAGGAGTTG	1021	DB	2851	TCCTGGCTCTTGGCGCGTATTTGCCGTCAACAGGCTGCGTGGTGTATAGTGGGAGGG
				QY	1982	TTGTCTCTGTCCGGAAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTGG
				DB	2911	TGCTCTTGTCCGGGAAGCGGCAATCATCTGACAGGGAAGTCTCTTACCGGAGAGTTGG
				QY	2042	ATGAAATGGAAGAGTGTCT 2059

Db 2971 ATGAGATGGAAGAGTGCT 2988

RESULT 8

109328 LOCUS 109328 5360 bp DNA linear PAT 02-DEC-1994

DEFINITION Sequence 8 from Patent WO 8904669.

ACCESSION 109328

VERSION 109328.1 GI:587963

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 5360)

AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.

JOURNAL Patent: WO 8904669-A 8 01-JUN-1989;

FEATURES

Location/Qualifiers

1..5360

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 86.7%; Score 1786; DB 6; Length 5360;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGCGCGCTATCAGCGCTATGCCAGCAGACAGGGCGCTTTGGGATGCATAATCACCA 61

Db 931 TGGCGCCCATCAGCGGTACGCCAGCAGACAGGGCGCTTCTAGGGTGCATATCACCA 990

QY 62 GCTTGACCGCGCGGCAAAAACAGGTGGAGGTTGAGTTTCAGATCGTGTCAATGCTG 121

Db 991 GCCTAACTGCGCGGACAAAACCAAGTGGAGGTTGAGTCCAGATTGTCACTGCTG 1050

QY 122 CCAGACTTTCTTGGCAACCTGCATTAAACCGGCTGTGGAGTGTCTACCATGGAGCG 181

Db 1051 CCCAAACCTTCTTGGCAACCTGCATTCAATGGGCTGTCTGAGTGTCTACCAACGGGCGG 1110

QY 182 GAACAAGGACATTGGCTCACCCTAAGGCTCTCTATCCAGATGTACCAATGTGGACC 241

Db 1111 GAACGAGACCATCGGCTCACCAGGGTCTCTATCCAGATGTATCCATATGTAGACC 1170

QY 242 AAGACCTCTGAGGCTGGCGGCTCCCAAGGTTCCCGCTCATTAACACCATGCATTGGC 301

Db 1171 AAGACCTTGTGGCTGGCGGCTCCGCAAGGTAGCGCTCATTTGACACCTGCACTTGG 1230

QY 302 GCTCTCGGACCTTTACCTGGTCAAGGACACCGCATGTCTCTGTGGCGCGGACGG 361

Db 1231 GCTCTCGGACCTTTACCTGGTCAAGGACACCGCATGTCTCTCGGTCGCGCGCGG 1290

QY 362 GTCATGGCAGGGCAGCGCTTTTCGCGCGGCTATCTCTACTTTGAAAGGCTCCTCGG 421

Db 1291 GTGATAGCAGGGCAGCGCTGTCTGCGCGGCGCCATTTCTTACTTTGAAAGGCTCCTCGG 1350

QY 422 GAGCGCTCTGTCTGTGCGCGCAGGACATGCCGTAGGCATATTTAGAGCGCGGTTAGCA 481

Db 1351 GGGGTCCGCTGTGTGCGCGCGGGCAGCGCGTGGGCATATTTAGGGCGCGGTGCA 1410

QY 482 CCGGTGAGGTGGCTAAGCGCGGTGGAATTCATCCCGGTAGAGGCTTAGAGACAACATGA 541

Db 1411 CCGGTGAGGTGGCTAAGCGCGGTGGAATTTATCCCTGTGGAGAACCTTAGAGACAACATGA 1470

QY 542 GGTCCCGGTGTTCTAGACAATCTCTCCCGCAGCAGTCCCGAGAGCTACCAAGTGG 601

Db 1471 GGTCCCGGTGTTCTAGACAATCTCTCTCCAGTAGTCCCGAGAGCTTCCAGTGG 1530

QY 602 CCCACCTGCATGCTCCACCGCAGCGGTAAAGACACCAAGGTCCCGCGCGCATACGCAG 661

Db 1531 CTACCTCTCATGCTCCACAGCAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCAG 1590

QY 662 CTCAGGCTCAAGGTTCTGGTGTCTCAACCCCTCGGTTGTGTGCAACATGGGCTTTGGTG 721

Db 1591 CTCAGGCTCAAGGTTCTGGTGTCTCAACCCCTCTGTGTGTGCAACACTGGGCTTTGGTG 1650

QY 722 TTATCATGTCCAAGGCCCATGGGATTTGATCTCTAAATCATCAGGACTGGGGTGAGCAATTA 781

Db 1651 TTATCATGTCCAAGGGCTATGGGATCGATCTCTAAATCATCAGGACCGGGGTGAGAACAAATTA 1710

QY 782 CTACTGGCAGCCCGATCAGCTATTCCACTAGCGGAAGTTCTTCCCGACGGCGGGTGT 841

Db 1711 CCACTGGCAGCCCGATCAGCTATTCCACTAGCGGAAGTTCTTCCCGACGGCGGGTGT 1770

QY 842 CAGGGGTGCTTATGACATAATAATTTGTGACAGTGCACATCCACGGATGCAATCCA 901

Db 1771 CGGGGGCGCTTATGACATAATAATTTGTGACAGTGCACATCCACGGATGCCACATCCA 1830

QY 902 TCTTGGCATTTGGCACTGTCTTGACCAAGACAGACCGCGGGGGGAGACTGACTGTGTG 961

Db 1831 TCTTGGCATTTGGCACTGTCTTGACCAAGACAGACTGCGGGGGGAGACTGCTGTGTG 1890

QY 962 TCGCCACCGTACCCCTCGGGCTCGGTCACTGTGCCCATCTTACATTCGAGGAGTTG 1021

Db 1891 TCGCCACCGTACCCCTCGGGCTCGGTCACTGTGCCCATCTTACATTCGAGGAGTTG 1950

QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATTA 1081

Db 1951 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGTATCA 2010

QY 1082 AGGGGGGAGACATCTCATCTTCTTCTGCACTCAAAAGAAAGTGCAGAGCTCGCGCAA 1141

Db 2011 AGGGGGGAGACATCTCATCTTCTTCTGCACTCAAAAGAAAGTGCAGAGAACTCGCGCAA 2070

QY 1142 AACTGTGCGGTGGCGCTCAATGCGTGGCTTACTACCGCGGCTTGTATGTCTCGTCA 1201

Db 2071 AGTGTGCGCATTTGGGCATCAATGCGTGGCTTACTACCGCGGCTTGTATGTCTCGTCA 2130

QY 1202 TCCGACACAGTGTGAGCTGTGCTGCGCAACTGACGCCCTCATGACCGGCTTTACCG 1261

Db 2131 TCCGACACAGCGGATGTTGCTGCTGCGCAACGATGCCCTCATGACCGGCTATACCG 2190

QY 1262 GCGACTTCGATTCGGTGTAGTGCACACACGCTGTGTACCCAGACAGTTCGACTTCAGCC 1321

Db 2191 GCGACTTCGACTCGGTGTAGTGCACATACGCTGTGTACCCAGACAGTTCGACTTCAGCC 2250

QY 1322 TTGACCTTACCTTACCATTTAGACATCAAGCTTCCCGAGGATGCTGTCTCCGTACTC 1381

Db 2251 TTGACCTTACCTTACCATTTAGACATCAAGCTTCCCGAGGATGCTGTCTCCCGACTC 2310

QY 1382 AACGTCGGGTAGGACTGCGCAGAGGAGCAGGCAATACAGATTTGTGGCACCGGGGG 1441

Db 2311 AACGTCGGGTAGGACTGCGCAGGAGGAGCAGGCAATACAGATTTGTGGCACCGGGGG 2370

QY 1442 AGCGTCTCTTCTGGCATTTGTGACTCGTCTGTCTGTGAGTGTCTATGACGCGGGTTG 1501

Db 2371 AGCGCCCTTCGGCATTTTGGACTCGTCTGTCTGTGAGTGTCTATGACGAGGCTGTG 2430

QY 1502 CTTGGTATGAGCTTACGCGCGCGGAGACACAGTTAGGCTACGAGATACATGAAACACC 1561

Db 2431 CTTGGTATGAGCTACGCGCGCGGAGACTACAGTTAGGCTACGAGGCTACATGAAACACC 2490

QY 1562 CGGACTTCCCGTGTGCCAAGACCATTTGAAATTTTGGGAGGCGCTTTTACGGGTCTCA 1621

Db 2491 CGGGGTCTCCGCTGTGCCAAGACCATTTGAAATTTTGGGAGGCGCTTTTACAGGCTCA 2550

QY 1622 CCACATAGACGCCCATTTCTCTATCCAGACAAAGCAGAGTGGGGAAACCTTTCCCTATC 1681

Db 2551 CTCATATAGTCCCATTTCTATCCAGACAAAGCAGAGTGGGGAAACCTTTCTTACC 2610

QY 1682 TGTAGGCTACCAAGCCACCGTGTGCTAGAGCTCAAGCCCTTCCCGCTCTGGGACC 1741

Db 2611 TGTAGGCTACCAAGCCACCGTGTGCTAGAGCTCAAGCCCTTCCCGCTCTGGGACC 2670

QY 1742 AGATGTGGAGTGTCTTGTGCTCAAGCCACCTTCCATGGGCCCAACACCTCTGTAT 1801

Db 2671 AGATGTGGAGTGTCTTGTGCTCAAGCCACCTTCCATGGGCCCAACACCTCTGTAT 2730


```
Db 2704 CTTGGTATAGCTCACCGCCGCGAGACTACAGTTAGGCTTACGAGGTACATGAACACCC 2763
Qy 1562 CGGAGCTTCCCGTGTGCGCAAGACCATCTTGAAATTTTGGGAGGCGCTCTTTACGGGTCTCA 1621
Db 2764 CGGGCTTCCCGTGTGCGCAGGACCATCTGAAATTTTGGGAGGCGCTCTTTACAGGCTCA 2823
Qy 1622 CCCACATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAAAACTTCCCTATC 1681
Db 2824 CTATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCCCTTACC 2883
Qy 1682 TGTAGCGTACCAAGCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
Db 2884 TGTAGCGTACCAAGCACCGTGTGCGCTAGAGGCTCAAGCCCTCCCGCATGTGGGACC 2943
Qy 1742 AGATGTGGAAGTCTTGATCCGCTCTCAAGCCACCCCTCATAGGCGCCAAACCTCTCTAT 1801
Db 2944 AGATGTGGAAGTCTTGATTCGCTCTCAAGCCACCCCTCATAGGCGCCAAACCCCTCTAT 3003
Qy 1802 ATAGATGGGCGCTGTCCAGAAATGAAGTCAACCTGACGCAACCCAGTCAACAAATATCA 1861
Db 3004 ACAGATGGGCGCTGTTCAGAAATGAAGTCAACCTGACGCAACCCAGTCAACAAATATCA 3063
Qy 1862 TGACATGTATGTGGGCTGACCTGGAGTGTCTACGAGTACCTGGGTGCTCTGTGGGCGG 1921
Db 3064 TGACATGTATGTGGGCGGACCTGGAGTGTCTACGAGTACCTGGGTGCTCTGTGGGCGG 3123
Qy 1922 TTCTGGCTGTTTGGCGCGTATTTGCTATCCACAGCTGCGTGTGTCATAGTGGTGGG 1981
Db 3124 TCTGGCTGTTTGGCGCGTATTTGCTATCCACAGCTGCGTGTGTCATAGTGGGCGG 3183
Qy 1982 TTCTTGTTCGGAAAGCCGCAATCATACCCGACAGGAAAGTCTCTACCGGAGTTGCG 2041
Db 3184 TCGTCTTGTTCGGAAAGCCGCAATCATACCTGACAGGGAAGTCTCTACCGGAGTTGCG 3243
Qy 2042 ATGAATGGAAGTGTCT 2059
Db 3244 ATGAGATGGAAGTGTCT 3261

RESULT 10
AR118696
LOCUS AR118696 7310 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 74 from patent US 6150087.
ACCESSION AR118696
VERSION AR118696.1 GI:14100606
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7310)
AUTHORS Chien,D.Y.
TITLE NANBV diagnostics and vaccines
JOURNAL Patent: US 6150087-A 74 21-NOV-2000;
FEATURES
    Location/Qualifiers
        source
            1..7310
                /organism="unknown"
                /mol_type="unassigned DNA"

ORIGIN
Query Match 86.7%; Score 1786; DB 6; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGGGCTATCAGGCGCTATGCCAGCAGACAAAGGGCGCTTTGGGATGCATTAATCACCA 61
Db 1729 TGGGGCGCATATCAGGCGCTATGCCAGCAGACAAAGGGCGCTCTTAGGTGCATTAATCACCA 1788
Qy 62 GCTTGACCGCGCGGACAAAACAGGTGAGGGTGGGTTTCAAGTCTGTCACATGCTG 121
Db 1789 GCCTAACTGGCGGGACAAAACAAAGTGGAGGGTGAAGTGTGTCACATGCTG 1848
Qy 122 CCCAGACTTCTTGGCAACCTGCAATTAACGGGGTGTGTGAGTGTCTACCATGGAGCGG 181
```

```
Db 1849 CCCAAACCTTCTGGCAACGTCATCAATGGGGTGTGTGGACTGTCTACCAACGGGCGG 1908
Qy 182 GAACAAGGACCATTCGGTCACCTAAGGGTCTGTTATCCAGATGTACACCAATGTGGACC 241
Db 1909 GAACGAGGACCATTCGGTCACCAAGGGTCTGTCATCCAGATGTATACCAATGTAGACC 1968
Qy 242 AAGACCTCGTAGGCTGGGCGCTCCCAAGGTGCCGCTCATTAACAACATGACATTTGGG 301
Db 1969 AAGACCTTGGGCTGGGCGCTCCGCAAGGTAGCGCTCAITTGACACCCCTGCACATTTGG 2028
Qy 302 GCTCTCTCGACCTTTTACCTGCTCAGAGGACGCGGATGTCATTCCTGTGCGCCGACGGG 361
Db 2029 GCTCTCTCGACCTTTTACCTGCTCAGAGGACGCGGATGTCATTCCTGTGCGCCGCGG 2088
Qy 362 GTGATGGCAGGCGGACCTGCTTTTGGCCCGGCTATCTTCTTACTTTGAAAGGCTCTCTCG 421
Db 2089 GTGATGAGGCGGACCTGCTGTGCGCCCGGCTATTTCTTACTTGAAGGCTCTCTCG 2148
Qy 422 GAGGCGCTCTGCTGTGCCCGGACGACATGCCGTAGGCTATTTCAAGCGCGGCTATGCA 481
Db 2149 GGGGTCCGCTGTGTGTGCCCGGCGGACGCGCTGGGCTATTTAGGGCGCGGCTGTGCA 2208
Qy 482 CCGGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGAGCTTAGAGACACCATCA 541
Db 2209 CCGGTGGAGTGGCTAAGCGGTGGACTTTCATCCCTGTGGAGAACCTTAGAGACACCATCA 2268
Qy 542 GGTGCCCGGTGTCTCAGACAACTCTCCCGACAGCAGTGTCCCGACAGCTTCAAGTGG 601
Db 2269 GGTGCCCGGTGTCTCAGGATAACTCTCTCCACAGTGTGTGCCCGACAGCTTCCAGTGG 2328
Qy 602 CCCACCTGATCTCTCCACCGGCGAGGTAAAGACCAAGAGTCCCGCGGCGATACGACAG 661
Db 2329 CTCACCTCATGCTCCACAGGCGGCAAGACCAAGGTCCCGGCTGCATATGCAAG 2388
Qy 662 CTCAGGCGTACAAGGTGTGTGTCTCAACCCCTCGTTCGTGTGCAACAATGTGGGCTTTGGTG 721
Db 2389 CTCAGGCGTATAAGGTGTGTGTCTCAACCCCTCTGTGTCTGCAACACTTGGGCTTTGGTG 2448
Qy 722 CTTACATGTCCAAAGCCCATGGATGTATCTTAAACATCAGGACTGGGGTGAAGACAATTA 781
Db 2449 CTTACATGTCCAAAGGCTCATGGATGTATCTTAAACATCAGGACTGGGGTGAAGACAATTA 2508
Qy 782 CTACTGGCAGCCCGATCAGTATTTCCACCTACGCGAAGTTCCTTGCAGCGCGGCTGT 841
Db 2509 CCACCTGGCAGCCCGATCAGTATTTCCACCTACGCGAAGTTCCTTGCAGCGCGGCTGT 2568
Qy 842 CAGGGGTGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGAGTGCACAACTCCA 901
Db 2569 CGGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGAGTGCACAACTCCA 2628
Qy 902 TCTTGGGCAATTGGCACTGTCTTGACCAAGCAGAGACCGGGGGGCGAGACTGACTGTGC 961
Db 2629 TCTTGGGCAATCGGCACTGTCTTGACCAAGCAGAGACTCGGGGGGCGAGACTGACTGTGC 2688
Qy 962 TCGCCACCGCTACCCCTCCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGTTG 1021
Db 2689 TCGCCACCGCTACCCCTCCGGGCTCGTCACTGTGCCCATCTCCCAACATCGAGGAGTTG 2748
Qy 1022 CTCTGTCCAATAACGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAAGCAATTA 1081
Db 2749 CTCTGTCCAACCAACGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAAGCAATTA 2808
Qy 1082 AGGGGGGAGACATCTCTCTTCTGCACTCAAGAGAGTGGCGAGCTGCGCGCA 1141
Db 2809 AGGGGGGAGACATCTCTCTTCTGCACTCAAGAGAGTGGCGAGCTGCGCGCA 2868
Qy 1142 AACTGTGCGGTGGGCGTCAATGCGGTGCTTACTACCGCGCTTGTGTGTCGTCA 1201
Db 2869 AGCTGTGCGATTTGGGATCAATGCGGTGCGCTTACTACCGCGCTTGTGTGTCGTCA 2928
Qy 1202 TCCCGACCAAGTGGTGTGTGTGTGCGCAACTGACGCCCTCATGACCGGCTTTTACC 1261
```

Db 2929 TCCGACACAGCGCGATGTTGTCGTGTCGCAACCGATGCCCTCATGACCGGTATACCG 2988
Qy 1262 GCGACTTCGATTCGGGTAGACTGCAACACGCTGTGTACCCAGACAGTGCATTCAGCC 1321
Db 2989 GCGACTTCGATTCGGGTAGACTGCAACACGCTGTGTACCCAGACAGTGCATTCAGCC 3048
Qy 1322 TTGACCTACCTTCACCATTTAGACAAATCACGCTTCCCGAGGATGCTGCTCCCGTACTC 1381
Db 3049 TTGACCTACCTTCACCATTTAGACAAATCACGCTTCCCGAGGATGCTGCTCCCGTACTC 3108
Qy 1382 AACGTCGGGTAGACTGCGAGGAAGCCAGGCAATACAGATTTGTGGACCGGGG 1441
Db 3109 AACGTCGGGTAGACTGCGAGGGAAGCCAGGCAATACAGATTTGTGGACCGGGG 3168
Qy 1442 AGCGTCTTCTGSCATGTTGACTGCTGCTGCTCTGCGAGTGTATGACGCGGTTGTG 1501
Db 3169 AGCGCCCTTCGCGATGTTGACTGCTGCTGCTCTGCGAGTGTATGACGCGGTTGTG 3228
Qy 1502 CTTGTATGAGCTTACGCCCGCGAGACCAACAGTTAGGCTACGAGCATACATGAACACC 1561
Db 3229 CTTGTATGAGCTTACGCCCGCGAGACTACAGTTAGGCTACGAGCATACATGAACACC 3288
Qy 1562 CGGACTTCGCTGTGCGCAAGACCATCTTGAATTTTGGAGGCGTCTTTACGGGTCTCA 1621
Db 3289 CGGAGCTTCCCGTGTGCGCAAGACCATCTTGAATTTTGGAGGCGTCTTTACAGGCTCA 3348
Qy 1622 CCACATAGACGCCCACTTCTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
Db 3349 CTCATATAGTGCCTTCTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 3408
Qy 1682 TGGTAGCTTACCAAGCCACCGTGTGCGTGTAGAGTCAAGCCCTTCCCGCTGTGGGACC 1741
Db 3409 TGGTAGCTTACCAAGCCACCGTGTGCGTGTAGAGTCAAGCCCTTCCCGCTGTGGGACC 3468
Qy 1742 AGATGTGGAAGTCTGTGATTCGCTTCAAGCCACCTTCATGGGCAACACCTTCTGCTAT 1801
Db 3469 AGATGTGGAAGTCTGTGATTCGCTTCAAGCCACCTTCATGGGCAACACCTTCTGCTAT 3528
Qy 1802 ATAGACTGGGCGGTGTCCAGATGAAGTCAACCTGACGACCAAGTCAACCAAGTATATCA 1861
Db 3529 ACAGACTGGGCGGTGTTCAGATGAATCAACCTGACGACCAAGTCAACCAAGTATATCA 3588
Qy 1862 TGACATGATGTGCGGTGACCTGAGAGTGTGTCAAGATACCTTGGGTGTGTTGGCGG 1921
Db 3589 TGACATGATGTGCGGTGACCTGAGAGTGTGTCAAGATACCTTGGGTGTGTTGGCGG 3648
Qy 1922 TTTGCTGCTTTTGGCGGATTTGCTTATCCACAGGCTGCGTGTATAGTAGGATGA 1981
Db 3649 TCTTGCTGCTTTTGGCGGATTTGCTTATCCACAGGCTGCGTGTATAGTAGGATGA 3708
Qy 1982 TTTGCTTGTCCGAAAGCCGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTG 2041
Db 3709 TGTCTTGTCCGGAAGCCGCAATCATACCTTGAACAGGAAATCTCTTACCGAGATTG 3768
Qy 2042 ATGAATGGAAGAGTGCT 2059
Db 3769 ATGAGATGGAAGAGTGCT 3786

RESULT 11

I09331
LOCUS 7310 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 15 from Patent WO 890469.
ACCESSION I09331
VERSION I09331.1 GI:587966
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7310)
AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
JOURNAL Patent: WO 890469-A 15 01-JUN-1989;
FEATURES Location/Qualifiers

source 1..7310
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 86.7%; Score 1786; DB 6; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
Qy 2 TGGCGCCTATCAGCGCTATGCCAGCAGACAAAGGGGCTTTTGGATGCATAATCACCA 61
Db 1729 TGGCGCCTATCAGCGCTATGCCAGCAGACAAAGGGGCTTCTAGGGTGCATAATCACCA 1788
Qy 62 GCTTACACCGCGCGGACAAACACAGGTGGAGGGTGCAGATTCAGATTCGATCGTCAATGCTG 121
Db 1789 GCTTACACCGCGCGGACAAACACAGGTGGAGGGTGCAGATTCGATTCGATTCGATCGTCA 1848
Qy 122 CCCAGACTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGACTGTCTACCAATGAGAGCGG 181
Db 1849 CCCAGACTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGACTGTCTACCAATGAGAGCGG 1908
Qy 182 GAACAGGACCATTCGCTCAGCTAAGGGTCCCTGTTATCCAGATGTACACCAATGTGAGCC 241
Db 1909 GAACAGGACCATTCGCTCAGCTAAGGGTCCCTGTTATCCAGATGTACACCAATGTGAGCC 1968
Qy 242 AAGACCTCGTAGGCTGGCGCCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGGC 301
Db 1969 AAGACCTCGTAGGCTGGCGCCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGGC 2028
Qy 302 GCTCTCGGACCTTTACCTGCTCAGAGGACGCGCGATGTCTTCTGTGCGCCGACGCGG 361
Db 2029 GCTCTCGGACCTTTACCTGCTCAGAGGACGCGCGATGTCTTCTGTGCGCCGACGCGG 2088
Qy 362 GTGATGGCAGGGGACGCTGCTTTCGCGCCCGGCTATCTTCTTACCTTGAAGGCTCCTCGG 421
Db 2089 GTGATGGCAGGGGACGCTGCTTTCGCGCCCGGCTATCTTCTTACCTTGAAGGCTCCTCGG 2148
Qy 422 GAGGCGCTCTGTGTCGCCCGCAGGACATGCCGTAGGATATTTACAGAGCCGCGGTATGCA 481
Db 2149 GAGGCGCTCTGTGTCGCCCGCAGGACATGCCGTAGGATATTTACAGAGCCGCGGTATGCA 2208
Qy 482 CCGGTGAGTGGCTTAAGCGGTGGAATTCATCCCGGTAGAGAGCTTAGAGACACCATGA 541
Db 2209 CCGGTGAGTGGCTTAAGCGGTGGAATTCATCCCGGTAGAGAGCTTAGAGACACCATGA 2268
Qy 542 GGTCCCGCGTGTCTCAGACAACTCCTCCACACAGAGTGCCTCCAGAGCTTCCAGGTGG 601
Db 2269 GGTCCCGCGTGTCTCAGACAACTCCTCCACACAGAGTGCCTCCAGAGCTTCCAGGTGG 2328
Qy 602 CCCACCTGCATGCTCCACCGCGCAGCGGTAAAGACCAAGGTCCCGGCGCATACGCAG 661
Db 2329 CTCACCTGCATGCTCCACCGCGCAGCGGTAAAGACCAAGGTCCCGGCGCATACGCAG 2388
Qy 662 CTCAGGCTCAAGGTGTGTGTCAACCTTCCGTTGCTGCAACATTTGGGCTTTGGTG 721
Db 2389 CTCAGGCTCAAGGTGTGTGTCAACCTTCCGTTGCTGCAACATTTGGGCTTTGGTG 2448
Qy 722 CTTACATGTCAAGGCGCCATGAGATTGATCTTAACATCAGAGCTGCGGTGAGACAAATTA 781
Db 2449 CTTACATGTCAAGGCGCCATGAGATTGATCTTAACATCAGAGCTGCGGTGAGACAAATTA 2508
Qy 782 TCTACTGGCAGCCCGCATCAGTATTCACCTACGCAAGTTCCTTCCCGAGCGCGGTGTT 841
Db 2509 CACCTGGCAGCCCGCATCAGTATTCACCTACGCAAGTTCCTTCCCGAGCGCGGTGTT 2568
Qy 842 CAGGGGTGCTTATGACATTAATTTGTGAGAGTGCATCCACGAGTGAACATCCA 901
Db 2569 CAGGGGTGCTTATGACATTAATTTGTGAGAGTGCATCCACGAGTGAACATCCA 2628
Qy 902 TCTTGGCATTTGGCAGCTGTCTTGAACAGAGACCGCGGGGCGAGACTGACTGTGC 961
Db 2629 TCTTGGCATTTGGCAGCTGTCTTGAACAGAGACTGCGGGGCGAGACTGACTGTGC 2688

TRCFDSTVTESDIRTEAIIYQCCDLDPQARVAIKSLITERLYVGGPLTNSRGENCGYRR
CRASGVLTSSGNTLTCYIKARAACHRAAGLODCTMLVCGDDLWLCESAGVEDAASL
RAPEAMTEYSAPDPDPPOPEVDELELITSCSNVSVAHDGAKRYVYLTRDFTPLAR
AAWETARHPVNSWLNIIINFAPTLWARMILMTHFVSVLARDQLEQALDCBIYACY
SIEPLDLPIIQL"

ORIGIN

Query Match	86.7%;	Score 1786;	DB 14;	Length 7310;
Best Local Similarity	91.7%;	Pred. No. 0;		
Matches 1888;	Conservative 0;	Mismatches 170;	Indels 0;	Gaps 0;
QY	2	TGCGCGCTATACGGCCTATGCCAGCAGACAAGGGCCCTTTGGGATGCATAATCACCA	61	
DB	1729	TGCGCGCCATCAGCGGTACGCCAGCAGACAAGGGCCCTCTAGGGTCATAATCACCA	1788	
QY	62	GCTTGACCGCCGGGAGAAAACACAGGTGGAGGGTGAGTTGAGTCGTCACTGCTG	121	
DB	1789	GCTTAATCGCGCGGAGAAAACCAAGTGGAGGGTGAGTCCAGATTGTGTCAACTGCTG	1848	
QY	122	CCCAGACTTTCTTGGCAACTGCTAATACCGGGTGTGGGACTGTCTACCATGGAGCGG	181	
DB	1849	CCCAACCTTCTTGGCAAGTGCATCAATGGGGTGCTGGACTGTCTACCAAGGGCCG	1908	
QY	182	GAACAAGGACCAATGGCTCACCTAAGGGTCTGTATTCCAGATGTACCAATGTGGACC	241	
DB	1909	GAACGAGGACCATCGCGTCACCAAGGGTCTGTCAATCCAGATGTATACCAATGTAGACC	1968	
QY	242	AGACCTCTAGCTGGCCGCTCCCAAGGTGCCCGCTCATTAACCATGCACTTGGC	301	
DB	1969	AGACCTTGTGGCTTGGCCGCTCCGCAAGGTAGCCGCTCATTTGACACCCCTGCCTTGGC	2028	
QY	302	GCTCTCGGACCTTTACCTGGTCAAGGACGCGGATGTCTTCTGTGGCCGCGGCGG	361	
DB	2029	GCTCTCGGACCTTTACCTGGTCAAGGACGCGGATGTCTTCTGTGGCCGCGGCGG	2088	
QY	362	GTGATGGAGGGGACGCTCTTTCGCCCGGCTATCTTCTTACTTGAAGGCTCTCTGG	421	
DB	2089	GTGATAGCAGGGGACGCTCTTTCGCCCGGCTATCTTCTTACTTGAAGGCTCTCTGG	2148	
QY	422	GAGGCCCTCTGTGTCGCCCGGACGACATGCCGTAGGCAATTCAGAGCGCGGTATGCA	481	
DB	2149	GAGGCTCCGTGTGTGTCGCCCGGCGGACGCGGTGGGCATTTTAGGCGCGCGGTGCA	2208	
QY	482	CCCGTGAGTGGCTAAGGCGGTGGACTTCATCCCGTAGAGACTTTAGAGACAACCATGA	541	
DB	2209	CCCGTGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGAGAACCTTAGAGACAACCATGA	2268	
QY	542	GGTCCCGGTGTCTCAGACAACTCTCCCAACAGAGTGCCTCCAGAGCTACCAAGTGG	601	
DB	2269	GGTCCCGGTGTCTCAGGATACTCTCTCCACAGTAGTGCCTCCAGAGCTTCCAGGTGG	2328	
QY	602	CCCACTGCTATGCTCCACCGGCGGCTAAGAGCACCAAGTCCCGGCGGATACGAG	661	
DB	2329	CTCACCTCATGCTCCCAAGGCGGCAAAAGCACCAAGTCCCGGCTGCATATGCA	2388	
QY	662	CTCAGGGCTACAGGTGCTGTGCTCAACCCCTCCGTTCTGCTGCAACATGGGCTTGGTG	721	
DB	2389	CTCAGGGCTTAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTGGTG	2448	
QY	722	CTTACATGTCCAAAGGCCATGGATGTATCTTAACATCAGGACTGGGGTGAGGACAATTA	781	
DB	2449	CTTACATGTCCAAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA	2508	
QY	782	CTACTGGACCGCATCAGTATTCACCTAGCGCAAGTTCCTTCCGACCGGGGTGTT	841	
DB	2509	CCACTGGACCGCCCATCAGTACTCACCTAGCGCAAGTTCCTTCCGACCGGGGTGCT	2568	
QY	842	CAGGGGGTCTTATGACATAATAATTGTGACGAGTGGCCACTCCACGATGCAACATCCA	901	
DB	2569	CAGGGGGTCTTATGACATAATAATTGTGACGAGTGGCCACTCCACGATGCAACATCCA	2628	
QY	902	TCTTGGGCATTTGCACTGCTTTGACCAAGCAGACCGCGGGGCGAGACTGACTGTGC	961	

DB	2629	TCTTGGGCATCGGCACACTGCTCTTGACAAAGCAGAGACTGCGGGGCGAGACTGGTGTGTC	2688	
QY	962	TGCGCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCCCATCTTACATCGAGGAGTTG	1021	
DB	2689	TGCGCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCCCATCTTACATCGAGGAGTTG	2748	
QY	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTGAAGCAATTA	1081	
DB	2749	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTGAAGTATCA	2808	
QY	1082	AGGGGGGAGAGACATCTCATCTTCTGCCACTCAAAGAAGTGCAGAGCTGCGCGCAA	1141	
DB	2809	AGGGGGGAGAGACATCTCATCTTCTGCCACTCAAAGAAGTGCAGAGACTGCGCGCAA	2868	
QY	1142	AACTGGTCCGCTTGGCGCTCAATGCGGTGCTTACTACCGGCTCTGATGTGCTCGTCA	1201	
DB	2869	AGCTGGTCCGCTTGGCGCTCAATGCGGTGCTTACTACCGGCTCTTGAAGTGTGCTCA	2928	
QY	1202	TCCCGACCAAGTGGTGCATCGTTCGTGGGCAACTGAGCGCCCTCATGACCGGCTTTACCG	1261	
DB	2929	TCCCGACCAAGTGGTGCATCGTTCGTGGGCAACTGAGCGCCCTCATGACCGGCTTTACCG	2988	
QY	1262	GGGACTTCGATTCGCTGATAGACTGCAACACGCTGTGTCAACCAGACAGTCTGACGCC	1321	
DB	2989	GGGACTTCGATTCGCTGATAGACTGCAACACGCTGTGTCAACCAGACAGTCTGACGCC	3048	
QY	1322	TTGACCCCTACCTTCAACATTTGAGCAATCAAGCTTCCCGAGATGCTGTCTCCGCTACTC	1381	
DB	3049	TTGACCCCTACCTTCAACATTTGAGCAATCAAGCTTCCCGAGATGCTGTCTCCGCTACTC	3108	
QY	1382	AACTGCGGGGTAGGACTGGCAGAGGAAAGCCAGGCACTTACAGATTTGTGGCACCGGGG	1441	
DB	3109	AACTGCGGGGTAGGACTGGCAGAGGAAAGCCAGGCACTTACAGATTTGTGGCACCGGGG	3168	
QY	1442	AGCGTCTTCTGGGCAATTTGATCTGCTGTCTCTCTGAGTGTCTATGACGCGGTTGTG	1501	
DB	3169	AGCGTCTTCTGGGCAATTTGATCTGCTGTCTCTCTGAGTGTCTATGACGCGGTTGTG	3228	
QY	1502	CTTGTGATGAGCTTACGCGCGCGGAGACCAAGTGTAGGCTACGAGCTACATGAACACCC	1561	
DB	3229	CTTGTGATGAGCTTACGCGCGCGGAGACCAAGTGTAGGCTACGAGCTACATGAACACCC	3288	
QY	1562	CGGGACTTCCCGGTGTCACCAAGACCACTTTGAATTTTGGAGGGCGCTTTTACGGGTCTCA	1621	
DB	3289	CGGGGCTTCCCGGTGTCACCAAGACCACTTTGAATTTTGGAGGGCGCTTTTACAGGCTCA	3348	
QY	1622	CCCAATAGACGCGCACTTCTTATCCAGACAAAGAGAGTGGGAAACCTTCCCTATC	1681	
DB	3349	CTCATATAGATGCCCACTTTCTATCCAGACAAAGAGAGTGGGAGAACCTTCCCTTACC	3408	
QY	1682	TGGTAGCTTACCAAGCCACCGTGTGCGGTAGAGCTCAAGCCCTCCCGCTGCGGACC	1741	
DB	3409	TGGTAGCTTACCAAGCCACCGTGTGCGGTAGAGCTCAAGCCCTCCCGCTGCGGACC	3468	
QY	1742	AGATGTGGAAGTGTGCTTGAATCGCTTCAAGCCACCCCTCATGGGCAACACCTCTGCTAT	1801	
DB	3469	AGATGTGGAAGTGTGCTTGAATCGCTTCAAGCCACCCCTCATGGGCAACACCTCTGCTAT	3528	
QY	1802	ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGCAACAGTCAACAGTATATCA	1861	
DB	3529	ACAGACTGGGCGCTGTTCAGAAATGAATCAACCTGACGCAACAGTCAACAGTATATCA	3588	
QY	1862	TGACATGTATGTGCGCTGACCTGAGGTGCTCAACAGTACCTGCTGTGCTGTGGCGG	1921	
DB	3589	TGACATGTATGTGCGCGGACCTTGAAGGTGCTCAACAGTACCTGCTGTGCTGTGGCGG	3648	
QY	1922	TTCTGGCTGTCTTGGCGCGTATTGCTTATCCAGAGGTGCGGTGCTCATAGTAGGAGGA	1981	
DB	3649	TCCTGGCTGTCTTGGCGCGTATTGCTTATCCAGAGGTGCGGTGCTCATAGTAGGAGGA	3708	
QY	1982	TTGTCTTGTGCGGAAAGCGGCAATCATACCGGACAGGAGTCTCTTACCGGGAGTTG	2041	
DB	3709	TCGTCTTGTGCGGAAAGCGGCAATCATACCTGACAGGAGTCTCTCTACCGAGAGTTG	3768	

QY 2042 ATGAATGGAAGAGTGCT 2059
Db 3769 ATGAGATGGAAGAGTGCT 3786

RESULT 13
I08294
LOCUS I08294 9185 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent EP 0388232.
ACCESSION I08294
VERSION I08294.1 GI:588994
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9185)
AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.
TITLE NABV diagnostics and vaccines
JOURNAL Patent: EP 0388232-A1 1 19-SEP-1990;
FEATURES
source 1. 9185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 86.7%; Score 1786; DB 6; Length 9185;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCCTATCAGCGCTATGCCAGCAGACAAAGGGCCCTTTTGGGATGCATAATCACCA 61
Db 3396 TGGCGCCCATCAGCGGTACGCCAGCAGACAAAGGGCCCTTCTAGGGTGCAATAATCACCA 3455

QY 62 GCTTGACCGGCCGGGCAAAAACACAGGTGAGGGTGAGGTTGAGTCTGATCTGCTGCTG 121
Db 3456 GCCTAACTGGCGGGGCAAAAACAAAGTGGAGGGTGAGGTTGATGTCACACTGCTG 3515

QY 122 CCAGACATTTCTTGGCAACTGCTAAACGGGGTGTGGAGTGTCTACCATGGAGCGG 181
Db 3516 CCAGACATTTCTTGGCAACTGCTAAACGGGGTGTGGAGTGTCTACCATGGAGCGG 3575

QY 182 GAACAAGGACCATGCGGTCACTAAGGGTCTCTGTTATCCAGATGTACACCAATGTGGACC 241
Db 3576 GAACAAGGACCATGCGGTCACTAAGGGTCTCTGTTATCCAGATGTACACCAATGTAGACC 3635

QY 242 AGACCTCTGAGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGACCTGGG 301
Db 3636 AAGACCTTGTGGCTGGCGCGCTCCCAAGGTGAGCGGCTCATTTGACACCCCTGCACTTGGC 3695

QY 302 GCTCTCGGACCTTTACCTGGTCAAGGACGACCGCATGTCATTCCTGTGGCGGACGGG 361
Db 3696 GCTCTCGGACCTTTACCTGGTCAAGGACGACCGCATGTCATTCCTGTGGCGGACGGG 3755

QY 362 GTGATGCGAGGGGAGCGCTCTTTCGCCCGCGCTATCTCTTACTTTGAAAGGCTCCTCGG 421
Db 3756 GTGATGCGAGGGGAGCGCTCTTTCGCCCGCGCTATCTCTTACTTTGAAAGGCTCCTCGG 3815

QY 422 GAGGCCCTCTGCTGTGCCCGGAGGACATGCCGTAGGCATATTCAGAGCGCGGATGCA 481
Db 3816 GGGGTCCTGCTGTGTGCCCGGAGGACATGCCGTAGGCATATTTAGGGCGCGGCTGTGCA 3875

QY 482 CCCGTGGAGTGGCTAAGGGGTGGACTTTCATCCCGTAGAGCTTTAGAGACCAACATGA 541
Db 3876 CCCGTGGAGTGGCTAAGGGGTGGACTTTCATCCCGTAGAGCTTTAGAGACCAACATGA 3935

QY 542 GGTCTCCCGGTGTTCTCAGACAACTCTCTCCCAAGCAGTGGCCAGAGCTACCAAGTGG 601
Db 3936 GGTCTCCCGGTGTTCTCAGACAACTCTCTCCCAAGCAGTGGCCAGAGCTTCCAGTGG 3995

QY 602 CCCACTGTGATGCTCCCAAGCGGAGCGGTAAAGACCAAGAGTCCCGGCGCATACGAG 661
Db 3996 CTCACTCCATGCTCCCAAGCGGAGCGGTAAAGACCAAGAGTCCCGGCGCATATGCGAG 4055

QY 662 CTCAGGGCTACAAGGTGCTGTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 721
Db 4056 CTCAGGGCTATAAGGTGCTGTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 4115

QY 722 CTTATATGTCCAAGGCCCATGGGATGATCTTAACATCAGGACTGGGGTGAGACAAATTA 781
Db 4116 CTTATATGTCCAAGGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA 4175

QY 782 CTACTGGGACCGCCGATCAGTATTCACCTACGCAAGTTCCTTGCCGACGGCGGGTGT 841
Db 4176 CCACTGGCAGGCCCATCAGTACTCACTACGCAAGTTCCTTGCCGACGGCGGGTGT 4235

QY 842 CAGGGGGTCTTATGACATAATAATTTGTGACAGTGGCCATCTCCAGGATGCAACATCCA 901
Db 4236 CGGGGGGGCTTATGACATAATAATTTGTGACAGTGGCCATCTCCAGGATGCAACATCCA 4295

QY 902 TCTTGGGCATTTGGCACTGCTCTTGACCAAGCAGAGACCGGGGGGAGACTGACTGTGC 961
Db 4296 TCTTGGGCATTCGGCACTGCTCTTGACCAAGCAGAGACTCGGGGGGAGACTGCTGTGTGC 4355

QY 962 TCGCCACCGCTACCCCTCCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGTTG 1021
Db 4356 TCGCCACCGCCACCCCTCCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGTTG 4415

QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAAGCAATTA 1081
Db 4416 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAAGTAATCA 4475

QY 1082 AGGGGGGAGAGACATCTCATCTTCTGCACTCAAGAGAGTGGCAGAGCTGCCCGGAA 1141
Db 4476 AGGGGGGAGAGACATCTCATCTTCTGCACTCAAGAGAGTGGCAGAGCTGCCCGGAA 4535

QY 1142 AACTGTGCGTGGGGCTCAATGCGGTGCTTACTACCGGCGCTTGTGATGTGTCGTCA 1201
Db 4536 AGCTGTGCGCATTTGGGCATCAATGCGGTGCTTACTACCGGCGCTTGTGATGTGTCGTCA 4595

QY 1202 TCCCGACCAAGTGGTGTGCTGCTGGGCAACTGACGCCCTCATGACCGGCTTTTACCG 1261
Db 4596 TCCCGACCAAGCGCGATGTTGCTGCTGGCAACGATGCCCTCATGACCGGCTATACCG 4655

QY 1262 GGGACTTCGATTCGGGTAGACTGCAACACGCTGTGTACCCAGACAGTGCAGCTTCAGCC 1321
Db 4656 GGGACTTCGACTTCGGGTAGACTGCAATACGCTGTGTACCCAGACAGTGCAGTTTCAGCC 4715

QY 1322 TTGACCTTACCTTCACTTGAAGCAATCAAGCTTCCAGAGTGTGCTCCCGTACTC 1381
Db 4716 TTGACCTTACCTTCACTTGAAGCAATCAAGCTTCCAGAGTGTGCTCCCGTACTC 4775

QY 1382 AACGTGGGGTAGGACTGGCAGAGGGAAGCCAGGCAATCTACAGATTTGTGGCACCGGGG 1441
Db 4776 AACGTGGGGTAGGACTGGCAGAGGGAAGCCAGGCAATCTACAGATTTGTGGCACCGGGG 4835

QY 1442 AGCGTCTTCTGGCATGTTGATCTGCTGTCTCTGCGAGTGTCTATGACGGGGTGTG 1501
Db 4836 AGCGTCTTCTGGCATGTTGATCTGCTGTCTCTGCGAGTGTCTATGACGGGGTGTG 4895

QY 1502 CTTGGTATGAGCTTACGCCCGGAGACCAAGTGTAGGCTACGAGCATATCAACACACC 1561
Db 4896 CTTGGTATGAGCTTACGCCCGGAGACCAAGTGTAGGCTACGAGCATATCAACACACC 4955

QY 1562 CGGGACTTCCCGTGTGCCAAGACCAATCTTGAATTTTGGGAGGGCGTCTTTTACGGGTCTCA 1621
Db 4956 CGGGACTTCCCGTGTGCCAAGACCAATCTTGAATTTTGGGAGGGCGTCTTTTACGGGTCTCA 5015

QY 1622 CCCACATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGGAAACCTTCCCTATC 1681
Db 5016 CTTATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGGAAACCTTCTTATACC 5075

QY 1682 TGGTAGCTTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
Db 5076 TGGTAGCTTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 5135

Db 4296 TCTTGGGCAATCGGCACTGCTTCCTTGACCAAGCAGAGACTCGGGGGGAGAGACTGGTTGTGC 4355
 Qy 962 TCGCCACCGCTACCCCTCGGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGGTG 1021
 Db 4356 TCGCCACCGCCACCCCTCGGGGCTCGGTCACTGTGCCCATCCCAACATCGAGGAGGTG 4415
 Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAGCAATTA 1081
 Db 4416 CTCTGTCCACCAACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAGCAATTA 4475
 Qy 1082 AGGGGGGAGAGATCTCACTCTTCGCACTCAAGAGAGAGTGCAGAGCTCGCCGCAA 1141
 Db 4476 AGGGGGGAGAGATCTCACTCTTCGCACTCAAGAGAGAGTGCAGAGCTCGCCGCAA 4535
 Qy 1142 AACTGTCTCGGTGGGCGTCAATGCGGTGGCTTACTACCGGGCGCTTGATGTCCGTCA 1201
 Db 4536 AGCTGTCTCGCATTTGGGCATCAATGCGGTGGCTTACTACCGGGCGCTTGATGTCCGTCA 4595
 Qy 1202 TCCGACACAGTGGTGTGAGCTGTGCTGTGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261
 Db 4596 TCCGACACAGCGGCGATGTGTGCTGTGGCAACCGATGCCCTCATGACCGGCTATACCG 4655
 Qy 1262 GCGACTTCGATTCGGGTAGACTGCAACACGCTGTGTACCCAGACAGTCTGACTTCAGCC 1321
 Db 4656 GCGACTTCGACTCGGTGTAGACTGCAATACGCTGTGTACCCAGACAGTCTGACTTCAGCC 4715
 Qy 1322 TTGACCCCTACCTTACCATGTAGACAATCAACGCTTCCCGAGGATGTCTCCGTACTC 1381
 Db 4716 TTGACCCCTACCTTACCATGTAGACAATCAACGCTTCCCGAGGATGTCTCCCGCACTC 4775
 Qy 1382 AACGTCGGGTAGGACTGCGAGGAGGAGCCAGGCACTTACAGATTTGGGCACCGGGG 1441
 Db 4776 AACGTCGGGCGAGGACTGGGAGGAGGAGCCAGGCACTTACAGATTTGGGCACCGGGG 4835
 Qy 1442 AGCGTCTCTCTGCACTGTTGACTCGTCTCTCTCGAGTCTATGACGCGGTTGTG 1501
 Db 4836 AGCGCCCCCTCGGCATGTTGACTCGTCTCTCTCTGAGTCTATGACGAGGCTGTG 4895
 Qy 1502 CTTGGTATGAGCTTACGCCCGCGAGACCAAGTCTAGGCTACGAGCATACATGACACCC 1561
 Db 4896 CTTGGTATGAGCTTACGCCCGCGAGACTACAGTCTAGGCTACGAGCGTACATGAACACC 4955
 Qy 1562 CGGCACTTCCCGTGTGCAAGACCATCTTGAATTTTGGAGGCGCTTTTACGGGTCTCA 1621
 Db 4956 CGGGGCTTCCCGTGTGCGAGACCATCTTGAATTTTGGAGGCGCTTTTACGGGTCTCA 5015
 Qy 1622 CCCACATAGACGCCCACTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
 Db 5016 CTCTATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 5075
 Qy 1682 TGGTAGGTTACCAAGCCCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
 Db 5076 TGGTAGGTTACCAAGCCCGTGTGCGCTAGGCTCAAGCCCTCCCGCTCGTGGGACC 5135
 Qy 1742 AGATGTGGAAGTGTCTGATCCGTCTCAAGCCCACTTCCATGGGCCCAACCTCTCTCTAT 1801
 Db 5136 AGATGTGGAAGTGTCTGATCCGTCTCAAGCCCACTTCCATGGGCCCAACCTCTCTCTAT 5195
 Qy 1802 ATAGACTGGGCGGTGTCCAGAAATGAAGTCAACCTGACGCAACCCAGTCAACCAATATCA 1861
 Db 5196 ACAGACTGGGCGGTGTTCAGAAATGAAGTCAACCTGACGCAACCCAGTCAACCAATATCA 5255
 Qy 1862 TGACATGTATGCGGCTGAGCTGGAGTGTGCTACAGAGTACCTGGGTGCTGTTGGCGGG 1921
 Db 5256 TGACATGTATGCGGCGGCTGAGGTGTGCTACAGAGTACCTGGGTGCTGTTGGCGGG 5315
 Qy 1922 TTCTGGGTGCTTTGGCGCGCTATTTGCTATCCACAGGCTCGGTGGTCAATAGTAGGAGGA 1981
 Db 5316 TCCTGGGTGCTTTGGCGCGCTATTTGCTGTCAACAGGCTCGGTGGTCAATAGTAGGAGGA 5375
 Qy 1982 TTGTCTTGTCCGGAAGCGGCAATCATACCGACAGGGAAGTCTCTACCGGAGTTTCG 2041
 Db 5376 TCGTCTTGTCCGGAAGCGGCAATCATACCGACAGGGAAGTCTCTACCGGAGTTTCG 5435

Qy 2042 ATGAANTGGAAGAGTCT 2059
 Db 5436 ATGAGATGGAAGAGTCT 5453

Search completed: February 27, 2005, 20:54:59
 Job time : 8691 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 15:10:47 ; Search time 1053 Seconds
(without alignments)
11586.502 Million cell updates/sec

Title: US-09-930-591-1
Perfect score: 2061
Sequence: 1 atggcgccatcacggccta.....atgaaatggaagagtgtga 2061

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2061	100.0	2061	6 AAD34500	Aad34500 Hepatitis
2	2061	100.0	2061	6 AAD31767	Aad31767 Hepatitis
3	2061	100.0	2061	10 AAD60868	Aad60868 Hepatitis
4	2061	100.0	2061	12 ADG47658	Adg47658 HCV NS3/4
5	1787	86.7	6299	4 AAF93669	Aaf93669 HCV NS3/4
6	1786	86.7	7310	1 AAN92106	Aan92106 Combined
7	1786	86.7	7310	1 AAN90336	Aan90336 Composite
8	1786	86.7	7310	2 AAQ98221	Aaq98221 Hepatitis
9	1786	86.7	8316	3 AAA75296	Aaa75296 cDNA sequ
10	1786	86.7	9133	2 AAZ07656	Aaz07656 Nucleotid
11	1786	86.7	9185	2 AAQ10566	Aaq10566 Hepatitis
12	1786	86.7	9185	3 AAA75297	Aaa75297 Sense str
13	1786	86.7	9185	12 AAD35979	Adn35979 HCV cDNA
14	1786	86.7	9401	2 AAT12710	Aat12710 Hepatitis
15	1786	86.7	9401	2 AAT99981	Aat99981 HCV polyp
16	1786	86.7	9401	2 AAV09989	Aav09989 HCV polyp
17	1786	86.7	9401	6 AAD35043	Aad35043 Hepatitis
18	1786	86.7	9401	12 ADL23106	Adl23106 Hepatitis
19	1786	86.7	9401	13 ADR29358	Adr29358 Hepatitis
20	1784.4	86.6	6905	1 AAN92103	Aan92103 Combined

21	1784.4	86.6	8316	2 AAQ05955	Aaq05955 Hepatitis
22	1784.4	86.6	9185	2 AAQ05956	Aaq05956 Sense str
23	1784.4	86.6	9185	2 AAX26737	Aax26737 Nucleotid
24	1784.4	86.6	9185	2 ADF66068	Adf66068 Hepatitis
25	1784.4	86.6	9400	2 AAQ21744	Aaq21744 Compiled
26	1782.8	86.5	9185	2 AAX00459	Aax00459 Hepatitis
27	1782.6	86.5	5300	1 AAN92097	Aan92097 Hepatitis
28	1781.2	86.4	2058	6 AAD29795	Aad29795 HCV-1 NS3
29	1781.2	86.4	2058	10 ADC06768	Adc06768 HCV mutan
30	1781.2	86.4	2058	12 ADL66804	Adl66804 HCV NS3/4
31	1781.2	86.4	5360	1 AAN90327	Aan90327 Hepatitis
32	1781.2	86.4	9618	11 ADN33102	Adn33102 Hepatitis
33	1779.6	86.3	2058	6 ABK15344	Abk15344 Hepatitis
34	1779.6	86.3	2058	10 ABX14410	Abx14410 DNA encod
35	1776.4	86.2	8316	12 ADN35977	Adn35977 HCV cDNA
36	1774.8	86.1	12980	2 AAV59364	Aav59364 Hepatitis
37	1774.8	86.1	12980	8 ACA62469	ACA62469 DNA encod
38	1774.8	86.1	16622	3 AAZ36212	Aaz36212 Nucleotid
39	1773.2	86.0	9379	2 AAQ36209	Aaq36209 Composite
40	1773.2	86.0	9646	2 AAV59361	Aav59361 Hepatitis
41	1773.2	86.0	9646	6 ABK87285	Abk87285 cDNA enco
42	1773.2	86.0	9646	8 ACA62466	ACA62466 HCV H77 c
43	1773.2	86.0	12980	6 ABK87286	Abk87286 Hepatitis
44	1765.2	85.6	9518	5 AAD03778	Aad03778 Hepatitis
45	1765.2	85.6	9599	2 AAX24833	Aax24833 Infectiou

ALIGNMENTS

RESULT 1
AAD34500
ID AAD34500 standard; DNA; 2061 BP.
XX
AC AAD34500;
XX
DT 16-JUL-2002 (first entry)
XX
DE Hepatitis C virus NS3/4A protein encoding DNA.
XX
KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; gene; ds.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS l..2061
FT /tag= a
FT /product= "HCV NS3/4A protein"
XX
PN WO200214362-A2.
XX
PD 21-FEB-2002.
XX
PR 15-AUG-2001; 2001WO-IB001774.
XX
PR 17-AUG-2000; 2000US-0225767P.
XX
PR 29-AUG-2000; 2000US-0229175P.
XX
PR 03-NOV-2000; 2000US-00705547.
XX
(TRIP-) TRIPEP AB.
XX
PI Sallberg M;
XX
DR WPI; 2002-339446/37.
XX
DR P-PSDB; AAE21837.
XX
PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.
XX
PS Claim 1; Page 64-65; 90pp; English.
XX

CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are
CC useful for identifying the presence or absence of HCV in a subject. They are
CC useful for preparing a medicament used for treating or preventing HCV
CC infection. Sequences of the invention are also used as vaccines. The
CC present sequence is a DNA encoding HCV NS3/4A protein

Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

Query Match				100.0%;	Score 2061;	DB 6;	Length 2061;
Best Local Similarity				100.0%;	Pred. No. 0;		
Matches 2061;				Conservative	0;	Mismatches	0;
				Indels	0;	Gaps	0;
Qy	1	ATGGCGCCTATCACGGCCTATGCCAGCAGACAAGGGCCCTTTTGGATGCATAATCAC	60				
Db	1	ATGGCGCCTATCACGGCCTATGCCAGCAGACAAGGGCCCTTTTGGATGCATAATCAC	60				
Qy	61	AGCTTGACCGCGCGGGA	120				
Db	61	AGCTTGACCGCGCGGGA	120				
Qy	121	GCCAGACTTTCTTGCCAACTGCAATTAACGGGGTGTGTGGACTGTCTACCATGGAGCC	180				
Db	121	GCCAGACTTTCTTGCCAACTGCAATTAACGGGGTGTGTGGACTGTCTACCATGGAGCC	180				
Qy	181	GGNAACAAGACCAATTGGCTCACCTTAAGGCTCTGTATACAGATGTACCAATATGTGGAC	240				
Db	181	GGNAACAAGACCAATTGGCTCACCTTAAGGCTCTGTATACAGATGTACCAATATGTGGAC	240				
Qy	241	CAAGACCTCGTAGGCTGGCCCGCTCCCCAAGGTGCCCGCTCATTTAACACCATGCACCTTGC	300				
Db	241	CAAGACCTCGTAGGCTGGCCCGCTCCCCAAGGTGCCCGCTCATTTAACACCATGCACCTTGC	300				
Qy	301	GGTCTCTCGGACCTTTA	360				
Db	301	GGTCTCTCGGACCTTTA	360				
Qy	361	GGTGATGGCAGGGGAGCGCTGTTTCGCCCGCGCTATCTTACTTGAAGGCTCTCTCG	420				
Db	361	GGTGATGGCAGGGGAGCGCTGTTTCGCCCGCGCTATCTTACTTGAAGGCTCTCTCG	420				
Qy	421	GGAGGCGCTCTGCTGTGCCCGCAGGACA	480				
Db	421	GGAGGCGCTCTGCTGTGCCCGCAGGACA	480				
Qy	481	ACCGTGAGTGGCTTAAGGGGTGCACTTCATCCCGTAGAGCTTTAGAGACAACCAATG	540				
Db	481	ACCGTGAGTGGCTTAAGGGGTGCACTTCATCCCGTAGAGCTTTAGAGACAACCAATG	540				
Qy	541	AGGTCCCGGTGTTCTCAGACAATCTCTCCCAACAGAGTGCCCGAGCTTACCAAGTG	600				
Db	541	AGGTCCCGGTGTTCTCAGACAATCTCTCCCAACAGAGTGCCCGAGCTTACCAAGTG	600				
Qy	601	GCCACCTGTCATGCTCCCAACGGCAGCGGTAAAGAGCACCAAGTCCCGCGCGCATACGCA	660				
Db	601	GCCACCTGTCATGCTCCCAACGGCAGCGGTAAAGAGCACCAAGTCCCGCGCGCATACGCA	660				
Qy	661	GCTCAGGGCTACAAGGTGCTGGTGTCAACCCCTCGTGTCTGCATCAATGGGCTTTTGT	720				
Db	661	GCTCAGGGCTACAAGGTGCTGGTGTCAACCCCTCGTGTCTGCATCAATGGGCTTTTGT	720				
Qy	721	GCTTACATGTCACAGGCCATGGGATTTGATCCTTAACATCAGGACTGGGTGAGGACAATT	780				
Db	721	GCTTACATGTCACAGGCCATGGGATTTGATCCTTAACATCAGGACTGGGTGAGGACAATT	780				
Qy	781	ACTACTGGCAGCCCGATCACGTAATTTCAACCTTACGGCAAGTTCCTTGCACAGCGGGTGT	840				
Db	781	ACTACTGGCAGCCCGATCACGTAATTTCAACCTTACGGCAAGTTCCTTGCACAGCGGGTGT	840				
Qy	841	TACAGGGGTGCTTATGACATTAATTTGTGAAGAGTGCCACTTCCACGATGCAACATCC	900				
Db	841	TACAGGGGTGCTTATGACATTAATTTGTGAAGAGTGCCACTTCCACGATGCAACATCC	900				

Db 1981 ATTGTTCTTCGGAAGCGGCAATCATACCGACAGGAATCTCTTACCGGAGTTC 2040
QY 2041 GATGAATGGAAGTGTCTGA 2061
Db 2041 GATGAATGGAAGTGTCTGA 2061

RESULT 2
AAD31767

ID AAD31767 standard; DNA; 2061 BP.
XX AC AAD31767;
XX DT 18-JUN-2002 (first entry)
XX DE Hepatitis C virus (HCV) NS3/4A DNA coding region.
XX KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytotatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
KW ds.
XX OS Hepatitis C virus.

XX FH Key Location/Qualifiers
FT CDS 1..2061
FT FT /*tag= a
FT FT /product= "HCV NS3/4A protein"
PN WO200213855-A2.

XX PD 21-FEB-2002.
XX PF 15-AUG-2001; 2001WO-1B001808.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX PA (TRIP-) TRIPEP AB.
XX PI Sallberg M, Hultgren C;
XX DR WPI; 2002-241837/29.
XX DR P-PSDB; AAE19900.
XX PT Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.

XX PS Claim 1; Page 94-95; 120pp; English.
XX CC The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
XX CC The present sequence is HCV NS3/4A DNA coding region

SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
Query Match 100.0%; Score 2061; DB 6; Length 2061;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCCTATCAGCGCCTATGCCAGACAGACAAGGGGCTTTTGGGATGCATATCACC 60
Db 1 ATGGCGCCTATCAGCGCCTATGCCAGACAGACAAGGGGCTTTTGGGATGCATATCACC 60
QY 61 AGCTTACCGCGCGGACAAAACAGGTTGAGGGTTCAGATCGTGTCAACTGCT 120

Db 61 AGCTTGAACGGCGCGGACAAAACACAGGTGAGGGTGAGGTTTCAGATCGTGTCAACTGCT 120
QY 121 GCCCAGACTTTCTTGGCAACCTGCAATTAACCGGGTGTGTGGACTGTCTTACCATGAGCC 180
Db 121 GCCCAGACTTTCTTGGCAACCTGCAATTAACCGGGTGTGTGGACTGTCTTACCATGAGCC 180
QY 181 GGAACAAGGACCAATTGCGTCACTTAAGGGTCTGTATTCCAGATGTACCAATGTGGAC 240
Db 181 GGAACAAGGACCAATTGCGTCACTTAAGGGTCTGTATTCCAGATGTACCAATGTGGAC 240
QY 241 CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCCCGCTATTAAACCATGCACTTGC 300
Db 241 CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCCCGCTATTAAACCATGCACTTGC 300
QY 301 GGCTCTCGGACCTTTTACCTGTCACGAGGCACGCGGATGTCAATTCCTGTGCGCCGACGG 360
Db 301 GGCTCTCGGACCTTTTACCTGTCACGAGGCACGCGGATGTCAATTCCTGTGCGCCGACGG 360
QY 361 GGTGATGGCAGGGGAGCGTGTTCGCGCCGCGCTATCTCTTACTTGAAGGCTCTCG 420
Db 361 GGTGATGGCAGGGGAGCGTGTTCGCGCCGCGCTATCTCTTACTTGAAGGCTCTCG 420
QY 421 GGAGGCCCTCTGCTGTGCCCCCAGGACATGCCGTAGGCATATTTCAGAGCCCGGATGC 480
Db 421 GGAGGCCCTCTGCTGTGCCCCCAGGACATGCCGTAGGCATATTTCAGAGCCCGGATGC 480
QY 481 ACCCGTGGAGTGGCTAAGCGGTGAGCTTCATCCCGTAGAGAGCTTAGAGACAACCATG 540
Db 481 ACCCGTGGAGTGGCTAAGCGGTGAGCTTCATCCCGTAGAGAGCTTAGAGACAACCATG 540
QY 541 AGGTCCCCGGTGTCTCAGACAATCTCTCCCAACAGCAGTGTGCCCGCAGAGCTACCAAGTG 600
Db 541 AGGTCCCCGGTGTCTCAGACAATCTCTCCCAACAGCAGTGTGCCCGCAGAGCTACCAAGTG 600
QY 601 GCCCAGCTGCATGCTCCCAACAGCAGTGTAGAGACCAAGGTCCCGCCGACATACGCA 660
Db 601 GCCCAGCTGCATGCTCCCAACAGCAGTGTAGAGACCAAGGTCCCGCCGACATACGCA 660
QY 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTGTCTGCTGCAACATATGGGCTTGGT 720
Db 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTGTCTGCTGCAACATATGGGCTTGGT 720
QY 721 GCTTACATGTCCAAGGCCCATGGGATGTATCTCAATCAGAGCTGGGGTGAGGACAAT 780
Db 721 GCTTACATGTCCAAGGCCCATGGGATGTATCTCAATCAGAGCTGGGGTGAGGACAAT 780
QY 781 ACTACTGGCAGCCCATCAGTATTTCCACCTAGCAGCAAGTTCCTTGCAGCGCGGCTGT 840
Db 781 ACTACTGGCAGCCCATCAGTATTTCCACCTAGCAGCAAGTTCCTTGCAGCGCGGCTGT 840
QY 841 TCAGGGGGTGTCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCAACATCC 900
Db 841 TCAGGGGGTGTCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCAACATCC 900
QY 901 ATCTTTGGCATTTGGCACTGTCTTCAACAGCAGAGACCGCGGGGCGAGACTGTGTG 960
Db 901 ATCTTTGGCATTTGGCACTGTCTTCAACAGCAGAGACCGCGGGGCGAGACTGTGTG 960
QY 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAAACATCGAGGAGTT 1020
Db 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAAACATCGAGGAGTT 1020
QY 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATATGCAAGGCTATTTCCTTCAAGCAAT 1080
Db 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATATGCAAGGCTATTTCCTTCAAGCAAT 1080
QY 1081 AAGGGGGGAGACATCTCATCTTCTGCGCACTCAAAAGAGAAGTGCAGCGAGCTGCGCGCA 1140
Db 1081 AAGGGGGGAGACATCTCATCTTCTGCGCACTCAAAAGAGAAGTGCAGCGAGCTGCGCGCA 1140
QY 1141 AAATGTTGCGGCTTGGGCGTCAATCGCGGCTTACTCCGCGGCTTGAATGTGTCCTG 1200
Db 1141 AAATGTTGCGGCTTGGGCGTCAATCGCGGCTTACTCCGCGGCTTGAATGTGTCCTG 1200


```
QY 421 GGAAGCCCTCTGCTGTGCCCGCAGGACATGCGGTAGGCATATTACAGAGCCGCGGTATGC 480
DB 421 GGAAGCCCTCTGCTGTGCCCGCAGGACATGCGGTAGGCATATTACAGAGCCGCGGTATGC 480
QY 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGAGCTTAGAGCAACCATG 540
DB 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGAGCTTAGAGCAACCATG 540
QY 541 AGGTCCCGCGTGTCTCAGACAACTCCTCCCAACAGCAGTGGCCAGAGCTACCAAGTG 600
DB 541 AGGTCCCGCGTGTCTCAGACAACTCCTCCCAACAGCAGTGGCCAGAGCTACCAAGTG 600
QY 601 GCCACACTGATCTCCACCGCAGCGGTAAGAGCAACCAAGTCCCGGCCGATACGCA 660
DB 601 GCCACACTGATCTCCACCGCAGCGGTAAGAGCAACCAAGTCCCGGCCGATACGCA 660
QY 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTTCGTCGCAACATGGGCTTTGGT 720
DB 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTTCGTCGCAACATGGGCTTTGGT 720
QY 721 GCTTACATGTCGAAGGCCCATGGATTGATCTTAAATCAGAGCTGGGGTGAGCAAT 780
DB 721 GCTTACATGTCGAAGGCCCATGGATTGATCTTAAATCAGAGCTGGGGTGAGCAAT 780
QY 781 ACTACTGGCAGCCGATCAGTATTCCACCTACGGCAAGTTCCTTGCAGCGCGGGTGT 840
DB 781 ACTACTGGCAGCCGATCAGTATTCCACCTACGGCAAGTTCCTTGCAGCGCGGGTGT 840
QY 841 TCAGGGGCTCTTATGACATAAATTTGTGACAGAGTGCCACTCCAGGATGCAATCC 900
DB 841 TCAGGGGCTCTTATGACATAAATTTGTGACAGAGTGCCACTCCAGGATGCAATCC 900
QY 901 ATCTTGGGCAATTCGCACTGCTTGACCAAGCAGAGACCGCGGGCGAGCTGACTGTG 960
DB 901 ATCTTGGGCAATTCGCACTGCTTGACCAAGCAGAGACCGCGGGCGAGCTGACTGTG 960
QY 961 CTGCGCACCTACCCCTCCGGGCTCCGTCACCTGTGCCCATCTTAACTCGAGGAGT 1020
DB 961 CTGCGCACCTACCCCTCCGGGCTCCGTCACCTGTGCCCATCTTAACTCGAGGAGT 1020
QY 1021 GCTCTGTCTACACCGGAGATCCCTTTATGGCAAGGCTATTCCCTTTGAAGCAAT 1080
DB 1021 GCTCTGTCTACACCGGAGATCCCTTTATGGCAAGGCTATTCCCTTTGAAGCAAT 1080
QY 1081 AAGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAAAGTCCGACGAGCTCGCGCA 1140
DB 1081 AAGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAAAGTCCGACGAGCTCGCGCA 1140
QY 1141 AAATGCTCGGTGGCGTCAATGCGGTGGCTTACTACCGCGGCTTGTATGTGTCGTC 1200
DB 1141 AAATGCTCGGTGGCGTCAATGCGGTGGCTTACTACCGCGGCTTGTATGTGTCGTC 1200
QY 1201 ATCCGACAGTGTGATGCTGTGCTGGGCAATGACGCGCTCATGACCGGCTTACC 1260
DB 1201 ATCCGACAGTGTGATGCTGTGCTGGGCAATGACGCGCTCATGACCGGCTTACC 1260
QY 1261 GCGCACTTCGATTCGGTATGACTGCAACAGTGTGTCACCCAGACAGTTCGACTTCAGC 1320
DB 1261 GCGCACTTCGATTCGGTATGACTGCAACAGTGTGTCACCCAGACAGTTCGACTTCAGC 1320
QY 1321 CTTGACCTTACCTTACCATTTGAGACAAATCAACGCTTCCCGAGGATGCTCTCCCGTACT 1380
DB 1321 CTTGACCTTACCTTACCATTTGAGACAAATCAACGCTTCCCGAGGATGCTCTCCCGTACT 1380
QY 1381 CAACGTCCGGGTAGGACTGCGAGAGGGAAGCCAGGCATCTACAGATTGTGGACCCGGGG 1440
DB 1381 CAACGTCCGGGTAGGACTGCGAGAGGGAAGCCAGGCATCTACAGATTGTGGACCCGGGG 1440
QY 1441 GAGGCTCTTCTGGCATTTTGTACTGCTGCTCTGCGAGTGTCTATGACGCGGTTGT 1500
DB 1441 GAGGCTCTTCTGGCATTTTGTACTGCTGCTCTGCGAGTGTCTATGACGCGGTTGT 1500
```

```
QY 1501 GCTTGTATGAGTTACGCCCGGAGACCAAGTTAGGCTACGAGCATACATGAACACC 1560
DB 1501 GCTTGTATGAGTTACGCCCGGAGACCAAGTTAGGCTACGAGCATACATGAACACC 1560
QY 1561 CCGGGAATTCCTCGTGTGCAAGACCATCTTGAATTTTGGAGGGCGCTCTTACGGGTCTC 1620
DB 1561 CCGGGAATTCCTCGTGTGCAAGACCATCTTGAATTTTGGAGGGCGCTCTTACGGGTCTC 1620
QY 1621 ACCCATAGAGCCGACCTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTAT 1680
DB 1621 ACCCATAGAGCCGACCTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTAT 1680
QY 1681 CTGGTAGCTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGCTGGGAC 1740
DB 1681 CTGGTAGCTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGCTGGGAC 1740
QY 1741 CAGATGTGAAGTGTGATCCGCTCTCAAGCCACCTCCATGGGCCAACACCTCTGCTA 1800
DB 1741 CAGATGTGAAGTGTGATCCGCTCTCAAGCCACCTCCATGGGCCAACACCTCTGCTA 1800
QY 1801 TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAGTATATC 1860
DB 1801 TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAGTATATC 1860
QY 1861 ATGACATGATGTGCGCTGACCTGGAGTGTGTCACAGTACCTGGGTGCTGTTGGCGGC 1920
DB 1861 ATGACATGATGTGCGCTGACCTGGAGTGTGTCACAGTACCTGGGTGCTGTTGGCGGC 1920
QY 1921 GTTCTGGCTGCTTTGGCGCGGTATTGCCCTATCCACAGGCTGGTGTCTAGTAGTAGG 1980
DB 1921 GTTCTGGCTGCTTTGGCGCGGTATTGCCCTATCCACAGGCTGGTGTCTAGTAGTAGG 1980
QY 1981 ATTGTCTGTCCGGAAGCCGCAATCATACCCGACAGGAAAGTCTCTACCCGGAGTTC 2040
DB 1981 ATTGTCTGTCCGGAAGCCGCAATCATACCCGACAGGAAAGTCTCTACCCGGAGTTC 2040
QY 2041 GATGAATGGAAGAGTGTGA 2061
DB 2041 GATGAATGGAAGAGTGTGA 2061

RESULT 4
ADG47658
ID ADG47658 standard; DNA; 2061 BP.
XX
AC ADG47658;
XX
DT 11-MAR-2004 (first entry)
XX
HCV NS3/4A domain DNA.
XX
ds; gene; immunogen; hepatitis C virus; HCV infection; vaccine.
Hepatitis C virus.
XX
Key Location/Qualifiers
FH 1..2061
FT /*tag= a
FT /product= "NS3/4A domain"
XX
US2003206919-A1.
XX
06-NOV-2003.
XX
26-NOV-2002; 2002US-00307047.
XX
17-AUG-2000; 2000US-0225767P.
XX
29-AUG-2000; 2000US-0229175P.
XX
15-AUG-2001; 2001US-00929955.
XX
15-AUG-2001; 2001US-00930591.
XX
(SALL/) SALLBERG M.
XX
```

PI Sallberg M;
 XX WPI: 2004-051480/05.
 DR P-PSDB; AD647659.
 DR
 PT New purified or isolated nucleic acid useful for enhancing an immune
 PT response to a hepatitis C antigen comprises specific nucleotide sequences
 PT and the amino acid sequences.
 XX
 XX Example 1; SEQ ID NO 1; 83pp; English.
 XX
 CC The invention relates to a purified or isolated nucleic acid. The
 CC peptides are useful as immunogens for the treatment and prevention of
 CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
 CC The nucleic acid and the peptide enhance an immune response to a
 CC hepatitis C antigen and are potent immunogens. The present sequence is
 CC used in the exemplification of the invention.
 XX
 SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
 Query Match 100.0%; Score 2061; DB 12; Length 2061;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCGCCTATACAGGCTATGCCAGCAGACAAAGGGGCTTTGGGATGCATATCACC 60
 DB 1 ATGGCGCCTATACAGGCTATGCCAGCAGACAAAGGGGCTTTGGGATGCATATCACC 60
 QY 61 AGCTTACCGCGCGGACAAAACCAAGGTGGAGGTGAGGTTGAGATCGTCAACTGCT 120
 DB 61 AGCTTACCGCGCGGACAAAACCAAGGTGGAGGTGAGGTTGAGATCGTCAACTGCT 120
 QY 121 GCCCAGACTTCTTGGCAACCTGCAATTAACGGGGTGTGTTGACTGTCTACCATGAGGC 180
 DB 121 GCCCAGACTTCTTGGCAACCTGCAATTAACGGGGTGTGTTGACTGTCTACCATGAGGC 180
 QY 181 GGAACAAAGGACCAATTCGTCACCTAAGGGTCTCTTTATCCAGATGTACCAATGTGGAC 240
 DB 181 GGAACAAAGGACCAATTCGTCACCTAAGGGTCTCTTTATCCAGATGTACCAATGTGGAC 240
 QY 241 CAAGACCTGTAGGCTGGCGCGTCCCAAGGTGCGCGCTCAATTAACCATGCACTTGC 300
 DB 241 CAAGACCTGTAGGCTGGCGCGTCCCAAGGTGCGCGCTCAATTAACCATGCACTTGC 300
 QY 301 GGTCTCTCGACCTTTACCTGTGACGAGCAGCGGATGTCTTCTGTGCGCGCAGCG 360
 DB 301 GGTCTCTCGACCTTTACCTGTGACGAGCAGCGGATGTCTTCTGTGCGCGCAGCG 360
 QY 361 GGTGATGGCAGGGGAGCCTGCTTTCCGCGCGGCTATCTTACTTTGAAAGGCTCTCG 420
 DB 361 GGTGATGGCAGGGGAGCCTGCTTTCCGCGCGGCTATCTTACTTTGAAAGGCTCTCG 420
 QY 421 GGAGGCGCTCTGTGTGCGCGGAGACATGCGGTAGGCATATTCAGAGCGCGGTATGC 480
 DB 421 GGAGGCGCTCTGTGTGCGCGGAGACATGCGGTAGGCATATTCAGAGCGCGGTATGC 480
 QY 481 ACCCGTGGAGTGGCTAAGCGGTGAGCTTATCCCGGTAGAGCTTTAGAGCAACCATG 540
 DB 481 ACCCGTGGAGTGGCTAAGCGGTGAGCTTATCCCGGTAGAGCTTTAGAGCAACCATG 540
 QY 541 AGGTCCCGGCTGTCTCAGACAACTCTCCCAACAGCAGTGCCTCAGAGCTACCAAGTG 600
 DB 541 AGGTCCCGGCTGTCTCAGACAACTCTCCCAACAGCAGTGCCTCAGAGCTACCAAGTG 600
 QY 601 GCCCACTGATGCTCTCCACCGGAGGTGAGACCAAGTCCCGCGCGCATACGCA 660
 DB 601 GCCCACTGATGCTCTCCACCGGAGGTGAGACCAAGTCCCGCGCGCATACGCA 660
 QY 661 GCTCAGGGCTACAAAGTGTGTGCTCAACCCCTCGTTGCTGCAACAATGGCTTGGT 720
 DB 661 GCTCAGGGCTACAAAGTGTGTGCTCAACCCCTCGTTGCTGCAACAATGGCTTGGT 720
 QY 721 GCTTACATGTCCAAAGGCCCATGGGATGTGATCTTAACATCAGGACTGGGTGAGGCAATT 780

DB 721 GCTTACATGTCCAAAGGCCCATGGGATGTGATCTTAACATCAGGACTGGGTGAGGCAATT 780
 QY 781 ACTACTGGCAGCGCGATCAGCTATTCCACCTACCGCAAGTTCTTGGCGAGCGGGGTGT 840
 DB 781 ACTACTGGCAGCGCGATCAGCTATTCCACCTACCGCAAGTTCTTGGCGAGCGGGGTGT 840
 QY 841 TCAGGGGGTGTATGACATAAATTTGTGACAGAGTGCCTCCTCACCGATGCAACATCC 900
 DB 841 TCAGGGGGTGTATGACATAAATTTGTGACAGAGTGCCTCCTCACCGATGCAACATCC 900
 QY 901 ATCTTGGGCAATGGCACTGTCTTGACCAAGCAGAGACCGGGGGCGAGACTGACTGTG 960
 DB 901 ATCTTGGGCAATGGCACTGTCTTGACCAAGCAGAGACCGGGGGCGAGACTGACTGTG 960
 QY 961 CTGCGCACCGGTACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAACATCGAGGAGTT 1020
 DB 961 CTGCGCACCGGTACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAACATCGAGGAGTT 1020
 QY 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATT 1080
 DB 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATT 1080
 QY 1081 AAGGGGGGAGACATCTCATCTTGTGCCATCTCAAGAAAGAGTGCGAGCTCGCCGCA 1140
 DB 1081 AAGGGGGGAGACATCTCATCTTGTGCCATCTCAAGAAAGAGTGCGAGCTCGCCGCA 1140
 QY 1141 AAACTGTGCGGTGGGGTCAATCCGTCGTCTACTACCGGGCTTGTGATGTGTCGTC 1200
 DB 1141 AAACTGTGCGGTGGGGTCAATCCGTCGTCTACTACCGGGCTTGTGATGTGTCGTC 1200
 QY 1201 ATCCCGACAGTGTGTGACGTTGTCTGTGGCACTGACGCGCTCATGACCGGCTTACC 1260
 DB 1201 ATCCCGACAGTGTGTGACGTTGTCTGTGGCACTGACGCGCTCATGACCGGCTTACC 1260
 QY 1261 GCGGACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 DB 1261 GCGGACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 QY 1321 CTGACCTCTACTTCAACATGAGCAATCAGCTTCCCGAGGATGCTGTCTCCGCTACT 1380
 DB 1321 CTGACCTCTACTTCAACATGAGCAATCAGCTTCCCGAGGATGCTGTCTCCGCTACT 1380
 QY 1381 CAACGCTCGGGGTAGGACTGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
 DB 1381 CAACGCTCGGGGTAGGACTGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
 QY 1441 GAGCGCTCTTCTGGCATGTTTGTGATCTGCTCTGCGAGTGTATGACGCGGTTGT 1500
 DB 1441 GAGCGCTCTTCTGGCATGTTTGTGATCTGCTCTGCGAGTGTATGACGCGGTTGT 1500
 QY 1501 GCTTGTATGAGCTTACGCGCGGAGACCAAGTGTAGGCTACGAGCATACATGAACACC 1560
 DB 1501 GCTTGTATGAGCTTACGCGCGGAGACCAAGTGTAGGCTACGAGCATACATGAACACC 1560
 QY 1561 CCGGAGCTTCCGCTGTGCAAGACCATCTTGAATTTGGAGGGGCTCTTTACGCGGTCTC 1620
 DB 1561 CCGGAGCTTCCGCTGTGCAAGACCATCTTGAATTTGGAGGGGCTCTTTACGCGGTCTC 1620
 QY 1621 ACCCACTAGAGCGGCTTCTTATCCAGACAAAGCAGAGTGGGGGAAACCTTCCCTAT 1680
 DB 1621 ACCCACTAGAGCGGCTTCTTATCCAGACAAAGCAGAGTGGGGGAAACCTTCCCTAT 1680
 QY 1681 CTGTGTAGGCTTACCAAGCAGCGGTGCGGTAGAGTCAAGCCCTTCCCGCTCGTGGGAC 1740
 DB 1681 CTGTGTAGGCTTACCAAGCAGCGGTGCGGTAGAGTCAAGCCCTTCCCGCTCGTGGGAC 1740
 QY 1741 CAGATGTGGAAGTGTGATCCGCTCAAGCCCACTTCAAGCCCACTTCAAGCCCACTTCA 1800
 DB 1741 CAGATGTGGAAGTGTGATCCGCTCAAGCCCACTTCAAGCCCACTTCAAGCCCACTTCA 1800
 QY 1801 TATAGCTGGGGCTGTCCAGATGAAGTCAAGCTGAGCGCAGCCAGTCAAGATATATC 1860

Db 1801 TATAGACTGGCGCTGTCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATATC 1860
QY 1861 ATGACATGATGTCGGCTGACCTGGAGGTGTCACAGGATACCTGGGTGCTGCTGGCGGC 1920
Db 1861 ATGACATGATGTCGGCTGACCTGGAGGTGTCACAGGATACCTGGGTGCTGCTGGCGGC 1920
QY 1921 GTTCTGGCTGCTTTGGCCGGGTATTCGCTATCCACAGGCTGGTGTCTAGTAGTAGG 1980
Db 1921 GTTCTGGCTGCTTTGGCCGGGTATTCGCTATCCACAGGCTGGTGTCTAGTAGTAGG 1980
QY 1981 ATTGCTTTGTCGGAAGCCGGAATCATACCCGACAGGAAAGTCTCTACCGGGAGTTC 2040
Db 1981 ATTGCTTTGTCGGAAGCCGGAATCATACCCGACAGGAAAGTCTCTACCGGGAGTTC 2040
QY 2041 GATGAATGGAAGTGTCTGA 2061
Db 2041 GATGAATGGAAGTGTCTGA 2061

RESULT 5

AAF83669
ID AAF83669 standard; DNA; 6299 BP.

AC AAF83669;

XX 23-JUL-2001. (first entry)

XX HCV NS34A ORF comprising pCMV-NS34A nucleic acid sequence.

XX HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;
KW catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;
KW immunotherapy; NS34A; ds.

XX Synthetic.

OS Hepatitis C virus.

XX Key Location/Qualifiers

FT CDS 1990..4050

FT /*tag= a

FT /*product= "NS34A"

XX WO200138360-A2.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US032326.

XX 24-NOV-1999; 99US-0167502P.

XX (CHIR) CHIRON CORP.

XX Coit D, Medina-Selby A, Selby M, Houghton M;

XX WPI; 2001-343948/36.

DR P-PSDB; AAB62633.

XX Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as
PT a vaccine against HCV, comprises a polypeptide having a mutation that
PT functionally disrupts the catalytic domain of NS3.

XX Disclosure; Fig 9; 340pp; English.

XX The invention relates to an isolated mutant non-structural (NS) Hepatitis
CC C virus (HCV) polypeptide, comprising a polypeptide having a mutation in
CC the catalytic domain of NS3, where the mutation functionally disrupts the
CC catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a
CC and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide
CC and polynucleotide (preferably DNA or a plasmid) compositions can be used
CC in vaccines against HCV and as diagnostics. The antibodies raised against
CC these polypeptides can also be used as diagnostics, or for passive
CC immunotherapy. The antibodies are also useful for isolating and
CC identifying HCV particles. The present sequence represents the nucleic
CC acid sequence of the pCMV-NS34A comprising the ORF encoding a NS34A

CC polypeptide

XX Sequence 6299 BP; 1449 A; 1713 C; 1637 G; 1500 T; 0 U; 0 Other;
SQ Query Match 86.7%; Score 1787; DB 4; Length 6299;
Best Local Similarity 91.7%; Pred No. 0;
Matches 1889; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 1 ATGCGCCCTATCACCGCCCTATGCCAGCAGACAAAGGGCCCTTTGGGATGCAATATCACC 60
Db 1990 ATGGCGCCCATCACGGCTAGCCCCAGCAGACAAAGGGCCCTCTAGGGGTGCATATCACC 2049
QY 61 AGCTTGACCGCGCGGAGCAAAACACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCT 120
Db 2050 AGCCTAACTGGCCGGGACAAAACCAAGTGGAGGTGAGGTTCCAGATTTGTCAACTGCT 2109
QY 121 GCCCAGACTTTCTTGGCAACCTGCAATTAACGGGGGTGTGTGAGCTGTCTACCATGAGGC 180
Db 2110 GCCCAACCTTCTTGCAACGTGCAATCAATGGGGGTGTGTGAGCTGTCTACCAAGGGGC 2169
QY 181 GGAACAAGACCAATTCGGTCACTTAAGGTCTCTTATCCAGATGTACACCAATGTGGAC 240
Db 2170 GGAACGAGGACCATCGCGTCAACCAGGGTCTCTCATCCAGATGTATACCAATGTAGAC 2229
QY 241 CAAAGCTCTGTAGGTGGCCCGCTCCCAAGGTGCCCGCTCATTAACACATGCATTCG 300
Db 2230 CAAAGCTTTGTGGGCTGGCCCGCTTCGCAAGGTACCCGCTCATTTGACACCTCTGCACTTC 2289
QY 301 GGCCTCTCGGACCTTTTACCTGTTCAGAGGCACGGATGTCTATCTCTGCGCCGACCG 360
Db 2290 GGCCTCTCGGACCTTTTACCTGTTCAGAGGCACGGATGTCTATCTCTGCGCCGCGG 2349
QY 361 GGTGATGGCAGGGGAGCCCTGCTTTTGGCCCGCCCTATCTCTTACTTGAAGGCTCTCTCG 420
Db 2350 GGTGATAGCAGGGGAGCCCTGCTGTCGCCCGCCGCTTCTTACTTGAAGGCTCTCTCG 2409
QY 421 GGAGGCCCTCTGCTGTGCCCCCAGACATGCGGTAGGATATTCAGAGCCGCGGTATGC 480
Db 2410 GGGGGTCCGCTGTGTGCCCCCAGGCGACGCGCGTGGGCATATTTAGGGCCGCGGTGTC 2469
QY 481 ACCCGTGGAGTGGCTAAGCGGTGAGCTTTCATCCCGTAGAGAGCTTAGAGACAACATG 540
Db 2470 ACCCGTGGAGTGGCTAAGCGGTGAGCTTTCATCCCTGTGGAGAACCTTAGAGACAACATG 2529
QY 541 AGGTCCCCGGTGTTCCTCAGACAACCTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG 600
Db 2530 AGGTCCCCGGTGTTCCTCAGACAACCTCTCTCCACAGTAGTGCCTCCAGAGCTTCCAGTG 2589
QY 601 GCCCACTGCTATGCTCCCAACCGGCGAGCGGTAAAGACCAAGGTCCCGCGCGCATACGCA 660
Db 2590 GCTCACTCCATGCTCCCAACAGGCGAGCGGCAAAAGCACCAAGGTCCCGGTGCATATGCA 2649
QY 661 GCTCAGGGCTACAAGGTGCTGGTGTCAACCCCTCGTGTGCTGCAACATGGCTTTGCT 720
Db 2650 GCTCAGGGCTACAAGGTGCTGGTGTCAACCCCTCGTGTGCTGCAACATGGCTTTGCT 2709
QY 721 GCTTACATGCTCAAGGCCCATCGGATTCATCTTAACATCAGGACTGGGTGGAGCAAT 780
Db 2710 GCTTACATGCTCAAGGCCCATCGGATTCATCTTAACATCAGGACTGGGTGGAGCAAT 2769
QY 781 ACTACTGGCAGCCGATACAGTATTCACCTACGGCAAGTTCCTTCCGACGCGGGGTGT 840
Db 2770 ACCACTGGCAGCCGATACAGTATTCACCTACGGCAAGTTCCTTCCGACGCGGGGTGT 2829
QY 841 TCAGGGGTGCTTATGACATATATTTGTGAGGTGCCATCTCCACGATGCCAATCC 900
Db 2830 TCAGGGGTGCTTATGACATATATTTGTGAGGTGCCATCTCCACGATGCCAATCC 2889
QY 901 ATCTTGGGATTTGGCACTGCTCTTACCAAGCAGACACCGCGGGGCGAGACTGACTGTG 960
Db 2890 ATCTTGGGATTTGGCACTGCTCTTACCAAGCAGACACCGCGGGGCGAGACTGACTGTG 2949
QY 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCTTAACTACGAGAGGTT 1020

|||||
2950 CTCGCCACCGCCACCCCTCCGGCTCCGTCACTGTGCCCTCCCAACATCGAGAGGTT 3009
|||||
1021 GCTCTGTCTCACTACCGGAGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAATT 1080
|||||
3010 GCTCTGTCTCACTACCGGAGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAATT 3069
|||||
1081 AAGGGGGGAGAGATCTCATCTTCTGCCACTCAAGAAAGAGTGGCAGAGCTCGCCGCA 1140
|||||
3070 AAGGGGGGAGAGATCTCATCTTCTGCCACTCAAGAAAGAGTGGCAGAGCTCGCCGCA 3129
|||||
1141 AAATCGTTCGCGTGGCGCTCAATCGGTGGCTTACTACCGCGGCTTGAATGTCGCGTC 1200
|||||
3130 AAGCTGCTGCATTTGGGCATCAATGCGTGGCTACTACCGCGGTCTTGACGTGTCGCTC 3189
|||||
1201 ATCCCGACCAAGTGGTACGTTGTCGTGCGTGGCACTGACGCGCTCATGACCGGCTTACC 1260
|||||
3190 ATCCCGACCAAGTGGTACGTTGTCGTGCGTGGCACTGACGCGCTCATGACCGGCTTACC 3249
|||||
1261 GCGGACTTCGATTCGGTGTAGACTGCAACAGCTGTGTCAACAGAGTGGCAGAGCTTCAGC 1320
|||||
3250 GCGGACTTCGATTCGGTGTAGACTGCAACAGCTGTGTCAACAGAGTGGCAGAGCTTCAGC 3309
|||||
1321 CTTGACCTTCACTTCACTTACGAGCAATCAGCTTTCCTCCAGGATGCTGTCTCCGTAAT 1380
|||||
3310 CTTGACCTTCACTTCACTTACGAGCAATCAGCTTTCCTCCAGGATGCTGTCTCCGTAAT 3369
|||||
1381 CACGTCGGGTAGGATGGCAGAGGGAAGCCAGGATCTACAGATTTGTGGCAGCGGG 1440
|||||
3370 CACGTCGGGTAGGATGGCAGAGGGAAGCCAGGATCTACAGATTTGTGGCAGCGGG 3429
|||||
1441 GAGCGTCTTCTGGCATGTTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
|||||
3430 GAGCGTCTTCTGGCATGTTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3489
|||||
1501 GCTTGTATGATGATACGCGCGGAGACCAAGTGTAGGCTACGAGCATACATGAACACC 1560
|||||
3490 GCTTGTATGATGATACGCGCGGAGACCAAGTGTAGGCTACGAGCATACATGAACACC 3549
|||||
1561 CCGGATCTTCCGTGTGCAAGACCATCTTGAATTTGGGAGGCGCTTTTACGGGTCTC 1620
|||||
3550 CCGGATCTTCCGTGTGCAAGACCATCTTGAATTTGGGAGGCGCTTTTACGGGTCTC 3609
|||||
1621 ACCACATAGAGCCGCTTCTATCCAGACCAAGAGTGGGAAACCTTCCCTAT 1680
|||||
3610 ACTCATATAGTCCGCTTCTATCCAGACCAAGAGTGGGAAACCTTCCCTAT 3669
|||||
1681 CTGGTAGCGTACCAAGCCACCGTGTGCTAGAGCTCAAGCCCTCCCTCCCTCGTGGGAC 1740
|||||
3670 CTGGTAGCGTACCAAGCCACCGTGTGCTAGAGCTCAAGCCCTCCCTCCCTCGTGGGAC 3729
|||||
1741 CAGATGTGAAGTGTGCTGCTCAAGCCACCTTCCATGGGCAACACCTCTGCTA 1800
|||||
3730 CAGATGTGAAGTGTGCTGCTCAAGCCACCTTCCATGGGCAACACCTCTGCTA 3789
|||||
1801 TATAGACTGGGCTGTCCAGATGAAGTCAACCTGACGACCAAGTCAACCAAGTATATC 1860
|||||
3790 TACAGACTGGGCTGTCCAGATGAAGTCAACCTGACGACCAAGTCAACCAAGTATATC 3849
|||||
1861 ATGACATGATGTGCGTGTGCTGAGGTGCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1920
|||||
3850 ATGACATGATGTGCGTGTGCTGAGGTGCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 3909
|||||
1921 GTCTGCTGCTTGTGGCGGTATGCTTATCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
|||||
3910 GTCTGCTGCTTGTGGCGGTATGCTTATCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3969
|||||
1981 ATTGTCTTGTCCGGAAGCGGCAATCATACCGGAGGGAAGTCTCTTACCGGAGTTC 2040
|||||
3970 GTCTGCTTGTCCGGAAGCGGCAATCATACCGGAGGGAAGTCTCTTACCGGAGTTC 4029
|||||
2041 GATGAATGGAGAGTGTCT 2059
|||||

Db 4030 GATGAGATGGAAGAGTGCT 4048

RESULT 6
AAN92106
ID AAN92106 standard; DNA; 7310 BP.
XX
AC AAN92106;
XX
DT 09-SEP-2004 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX
DE Combined ORFs of the HCV cDNAs from clones K9-1 through 15e.
XX
KW Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH; ss.
XX
OS Hepatitis C virus.
OS Unidentified.
XX
FH Key
FT CDS Location/Qualifiers
EP318216-A.
XX
XX 31-MAY-1989.
XX
XX 18-NOV-1988; 88EP-00310922.
XX
XX 18-NOV-1987; 87US-00122714.
XX 30-DEC-1987; 87US-00139886.
XX 26-FEB-1988; 88US-00161072.
XX 06-MAY-1988; 88US-00191263.
XX 26-OCT-1988; 88US-00263584.
XX 14-NOV-1988; 88US-00271450.
XX
XX (CHIR) CHIRON CORP.
XX (CHIR) CHIRON CORP.
XX
XX Houghton M, Choo QL, Kuo G;
XX WPI; 1989-159274/22.
XX P-PSDB; AAP92050.
XX
XX Purified hepatitis C virus - and associated nucleic acids and
XX polypeptide(s).
XX
XX Claim 3; Fig 47-1 - 47-8; 139pp; English.
XX
XX It is a double-stranded nucleotide sequence of the open reading frame
XX (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus
XX (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to
XX detect the presence of HCV nucleic acids in samples. The polypeptide(s)
XX it encodes could be used as immunoassay reagents and vaccines and to
XX generate antibodies useful in diagnosis and passive immunotherapy for HCV
XX infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR
XX field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Revised record issued on 09-SEP-2004 : Correction to keywords
XX
XX Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 0 U; 4 Other;
SQ

Query Match 86.7%; Score 1786; DB 1; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCCCTATCACGCGCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCATAATCACCA 61
Db 1729 TGGCGCCCTATCACGCGCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCATAATCACCA 1788
QY 62 GCTTACCGCGCGGACAAAACAGGTGGAGGTTCAGATCGTGTCACTGCTG 121
|||

Db 1789 GCCTAACTGCGCGGGA CA AAAA CCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCTG 1848
Qy 122 CCCAGACTTTCTTTGGCAACCTGCAATAACCGGGGTGTGTTGCACTGTCTACCATGAGACCG 181
Db 1849 CCCAAACCTTCTTGGCAACGTGCATCAATGGGGTGTGCTGCACTGTCTACCAACGGGGCG 1908
Qy 182 GAACAGGACCAATTGCGTCACTTAAGGGTCTGTATTACAGATGTACCAATGTGGACC 241
Db 1909 GAACGAGGACCATCGCGTCAACCAAGGGTCTGTATCCAGATGTATACCAATGTAGACC 1968
Qy 242 AAGACCTCGTAGCTGCGCGCTCCCAAGGTGCGCTCATTAACACCATGACATTTGG 301
Db 1969 AAGACCTTGTGGCTGCGCGCTCCGCAAGTAGCGCTCATTAAGACACCTGACATTTGG 2028
Qy 302 GCTCTCGGACCTTTACTCTGTCACGAGGACGCGGATGTCAATTCCTGTGCGCGGACGG 361
Db 2029 GCTCTCGGACCTTTACTCTGTCACGAGGACGCGGATGTCAATTCCTGTGCGCGGCGG 2088
Qy 362 GTGATGCGAGGGCAGCTGCTTTGCGCGCGCTATCTCTTAATTGAAGGTCTCTCGG 421
Db 2089 GTGATAGCAGGGCAGCTGCTGTGCGCGCGGCTATTTAGGGCGCGCGGTGTGCA 2148
Qy 422 GAGGCTCTGCTGTGCGCGCGGACATGCGGTAGGCATATTCAGAGCGCGCGGTATGCA 481
Db 2149 GGGTCCGCTGTTGTGCGCGCGGGGCAACGCGTGGGCATATTTAGGGCGCGCGGTGTGCA 2208
Qy 482 CCGGTGAGTGGTCAAGCGGTGACATTCATCCCGTAGAGAGCTTAGAGACAACCATGA 541
Db 2209 CCGGTGAGTGGTCAAGCGGTGACATTCATCCCGTAGAGAGCTTAGAGACAACCATGA 2268
Qy 542 GGTGCGCGGTGTTCTCAGACAACCTCTCCCAAGCAGAGTCCCGCAGAGCTACCAAGTGG 601
Db 2269 GGTGCGCGGTGTTCAAGGATAAATCTCTCTCCACAGTAGTGCCCGCAGAGCTTCCAGTGG 2328
Qy 602 CCACCTGATGCTCCACCGCGAGCGGTAAAGAGCACCAAGTCCCGCGGATACGCAG 661
Db 2329 CTCACCTCCATGCTCCACAGGAGCAGCGGCAAAAGCACCAAGTCCCGGCTGCATATGCG 2388
Qy 662 CTCAGGCTCAAGGTGCTGCTCAACCCCTCCGTTGTGCAACAATGGGCTTTGGTG 721
Db 2389 CTCAGGCTCAAGGTGCTGCTCAACCCCTCCGTTGTGCAACAATGGGCTTTGGTG 2448
Qy 722 CTTCATGTCCTCAAGGCTCATGGATGATCTTAACATCAGGAGCTGGGGTGAGGACAATTA 781
Db 2449 CTTCATGTCCTCAAGGCTCATGGATGATCTTAACATCAGGAGCTGGGGTGAGGACAATTA 2508
Qy 782 CTACTGCGAGCGCATCAGTATTCACCTACGGCAAGTTCCTTGGCGAGCGGGGTGT 841
Db 2509 CCACTGCGAGCGCCCATCACGTACTTCCACCTACGGCAAGTTCCTTGGCGAGCGGGGTGT 2568
Qy 842 CAGGGGTGCTTATGACATAATATTGTGACAGTGCCCACTCCACGGATGCAACATCCA 901
Db 2569 CAGGGGTGCTTATGACATAATATTGTGACAGTGCCCACTCCACGGATGCAACATCCA 2628
Qy 902 TCTTGGCATTTGGCACTGTCTTGACCAAGCAGAGACCGCGGGGCGGAGACTGACTGTGC 961
Db 2629 TCTTGGCATTTGGCACTGTCTTGACCAAGCAGAGACTGCGGGGCGGAGACTGTTGTGC 2688
Qy 962 TCGCACCGTACCCCTCGGGTTCGTCATCTGTGCCCCATCTTAACATCAGAGAGTTG 1021
Db 2689 TCGCACCGTACCCCTCGGGTTCGTCATCTGTGCCCCATCTTAACATCAGAGAGTTG 2748
Qy 1022 CTCTGTCCACTACCGGAGATCCCTTTTATGGCAGGCTATTTCCCTTGAAGCAATTA 1081
Db 2749 CTCTGTCCACTACCGGAGATCCCTTTTATGGCAGGCTATTTCCCTTGAAGCAATTA 2808
Qy 1082 AGGGGGGAGACATCTCTCTTGTGCGCACTCAAGAGAGAGTGCAGAGCTCGCGCAA 1141
Db 2809 AGGGGGGAGACATCTCTCTTGTGCTATTCAGAGAGAGTGCAGAGCTCGCGCAA 2868
Qy 1142 AACTGTGCGGTGGGCGTCAATGCCGTGCTTACTACCGCGCTTGTATGTGTCGTC 1201
Db 2869 AGCTGGTTCGATTTGGGATCAATGCGCGTGTACTACCGCGTCTTGAAGTGTGTCGTC 2928

Qy 1202 TCCCGACCAAGTGTGAGCTGTTGTGTCGTGGCAACTGACGCGCTCATGACCGGCTTTACCG 1261
Db 2929 TCCCGACCAAGTGTGAGCTGTTGTGTCGTGGCAACGATGCCCTCATGACCGGCTATACCG 2988
Qy 1262 GCGACTTCATTCGTTGCTGATAGACTGCAACACGTTGTGTACCCAGACAGTGCATTCAGCC 1321
Db 2989 GCGACTTCGACTCGGTGATAGACTGCAATACGTTGTGTACCCAGACAGTGCATTCAGCC 3048
Qy 1322 TTGACCTTACCTTTACATTTGAGACAATCAACGCTTCCCGAGGATGCTGTCTCCGCTACTC 1381
Db 3049 TTGACCTTACCTTTACATTTGAGACAATCAACGCTTCCCGAGGATGCTGTCTCCGCTACTC 3108
Qy 1382 AACGTCGGGGTAGGACTGCGCAGAGGGAAGCCAGGCAATTCACAGATTGTGGCACCGGGGG 1441
Db 3109 AACGTCGGGGTAGGACTGCGCAGGGAAGCCAGGCAATTCACAGATTGTGGCACCGGGGG 3168
Qy 1442 AGGTCCTTCTGCGCATGTTTGTGCTGTCTGTCTGTGCGAGTGTCTATGACGCGGGTGTG 1501
Db 3169 AGCGCCCTTCCGCGCATGTTTGTGCTGTCTGTCTGTGAGTGTCTATGACGAGGCTGTG 3228
Qy 1502 CTTGGTATGAGCTTACGCGCGCGGAGCACTTCTTAAATTTTGGGAGGCGCTCTTTACGGGTCTCA 1561
Db 3229 CTTGGTATGAGCTTACGCGCGCGGAGCACTTCTTAAATTTTGGGAGGCGCTCTTTACGGGTCTCA 3288
Qy 1562 CCGGACTTCCCGGTGTGCGCAAGACCACTTCTTAAATTTTGGGAGGCGCTCTTTACGGGTCTCA 1621
Db 3289 CCGGACTTCCCGGTGTGCGCAAGACCACTTCTTAAATTTTGGGAGGCGCTCTTTACGGGTCTCA 3348
Qy 1622 CCGACATAGACGCCCATCTTCTATCCAGACAAGAGAGTGGGGGAAACCTTTCCTTATC 1681
Db 3349 CTCATATAGATGCCCACTTCTATCCAGACAAGAGAGTGGGGGAAACCTTTCCTTATC 3408
Qy 1682 TGGTAGCTACCAAGCCCGGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGTGGGAGC 1741
Db 3409 TGGTAGCTACCAAGCCCGGTGTGCGCTAGAGCTCAAGCCCTCCCGCTGTGGGAGC 3468
Qy 1742 AGATGTGGAAGTGTCTGATCTCAAGCCCAACCTCCATGGGCGCAACACCTCTGCTAT 1801
Db 3469 AGATGTGGAAGTGTCTGATCTCAAGCCCAACCTCCATGGGCGCAACACCTCTGCTAT 3528
Qy 1802 ATAGACTGGGCTGTGTCAGAAATGAAGTCAACCTGAGCAACCCAGTCAACAAATATCA 1861
Db 3529 ACAGACTGGGCTGTGTCAGAAATGAAGTCAACCTGAGCAACCCAGTCAACAAATATCA 3588
Qy 1862 TGACATGTATGTGCGCTGACCTGAGGTGTGTCAGAGTACCTGGGTGCTGTTGGCGGG 1921
Db 3589 TGACATGTATGTGCGGCGGACCTGGAGTGTGTCAGAGCAACCTGGGTGCTGTTGGCGGG 3648
Qy 1922 TTCTGGCTCTTTTGGCGGCTATTTGCTATCCACAGGCTGCGTGGTATAGTAGGTAGGA 1981
Db 3649 TTCTGGCTCTTTTGGCGGCTATTTGCTATCCACAGGCTGCGTGGTATAGTAGGTAGGA 3708
Qy 1982 TTGTCTTGTCCGGAAGCGGCAATCATACCGAGAGGAAAGTCTTCTACCGGAGTTCG 2041
Db 3709 TTGTCTTGTCCGGAAGCGGCAATCATACCTGACAGGAAAGTCTTCTACCGAGAGTTCG 3768
Qy 2042 ATGAAATGCAAGTGTCT 2059
Db 3769 ATGAGATGGAAGTGTCT 3786

RESULT 7

AAN90336

ID AAN90336 standard; DNA; 7310 BP.

XX

AAN90336;

XX

25-MAR-2003 (revised)

DT

19-JUL-2001 (revised)

DT

01-NOV-1989 (first entry)

XX

DE Composite hepatitis C virus (HCV) cDNA.

XX Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.
KW Pan troglodytes.
OS GB2212511-A.
PN 26-JUL-1989.
PD 18-NOV-1988; 88GB-00027024.
XX 18-NOV-1987; 87US-00122714.
XX 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 26-OCT-1988; 88US-00263584.
XX (CHIR) CHIRON CORP.
PA Houghton M, Choo QL, Kuo G;
XX WPI; 1989-215054/30.
XX P-PSDB; AAP90288.
DR Hepatitis C virus gene - used for prodn. of polynucleotide probes
XX polypeptide(s) and antibodies for diagnosis, prevention and treatment of
XX infection.
XX Disclosure; Fig 47; 30pp; English.
XX The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by
CC aligning clones k9-1 through 15e in 5'-3' direction. The cDNA encodes
CC antigens which react with antibodies in patients with non-A non-B
CC hepatitis (NANBH). The cDNA can be used to design probes, or to
CC synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to
CC raise antibodies for immunoassay or treatment, or to produce vaccines.
CC See also AAP90288, and AAN90303-35. (N.B. This record was resubmitted to
CC correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR
XX field.)
XX Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 U; 0 Other;
SQ
Query Match 86.7%; Score 1786; DB 1; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 2 TGCGCGCTATCAGCGCTATGCCAGCAGACAAAGGGCCCTTTGGATGCATAATCACCA 61
DB 1729 TGCGCGCCCATCAGCGGTACGCCAGCAGACAAAGGGCCCTCCTAGGGTGCATAATCACCA 1788
QY 62 GCTTGACCGCGGGACAAAACAGGTGGAGGTGAGGTTCAGATCGTCAACTGCTG 121
DB 1789 GCTTAAGTGGCGGGGACAAAACAGGTGGAGGTGAGGTTCAGATCGTCAACTGCTG 1848
QY 122 CCAGACTTTCTTGGCAACCTGCAATTAACGGGTGTGTTGGACTGTCTACCATGGAGCG 181
DB 1849 CCCAAACCTTCTGGCAACCTGCAATTAACGGGTGTGTTGGACTGTCTACCATGGAGCG 1908
QY 182 GAACAAGGACCATTTGGTCACTAAGGTCTGTTATCAGATGTACCAATGTGGACC 241
DB 1909 GAACGAGGACCATTCGGCTCACCAGGGTCTGTCATCCAGATGTATACCAATGTAGACC 1968
QY 242 AAGACCTGTAGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGG 301
DB 1969 AAGACCTGTGGGCTGGCGCGCTCCCAAGGTAGCGCTCATTAACACCATGCACTTGG 2028
QY 302 GCTCTCGGACCTTTACCTGGTCAAGGACAGCGCGATGTCAATCTGTGCGCGGACGG 361
DB 2029 GCTCTCGGACCTTTACCTGGTCAAGGACAGCGCGATGTCAATCTGTGCGCGGACGG 2088
QY 362 GTGATGAGGGGACGCTGCTTTCGCCCGGCTATCTTACTTTGAAGGCTCTCTCG 421
DB 2089 GTGATGAGGGGACGCTGCTGTCGCCCGGCTATCTTCTACTTTGAAGGCTCTCTCG 2148

QY 422 GAGGCCCTCTGCTGTGCCCGCAGGACATGCCGTAGGCATATTCAGAGCCGCGGTATGCA 481
DB 2149 GGGGTCCGCTGTGTGTCGCCCGGCGCAGCGTGGGCATATTTAGGCGCGGCTGTGCA 2208
QY 482 CCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGAGCTTTAGAGACAACATGA 541
DB 2209 CCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGAACCTAGAGACAACATGA 2268
QY 542 GGTCCCGGTGTCTCAGACAACCTCTCCCAACAGCAGTGTCCCGAGAGCTACCAAGTG 601
DB 2269 GGTCCCGGTGTCTCAGACAACCTCTCCCAACAGTGTCCCGAGAGCTTTCCAGGTGG 2328
QY 602 CCCACTGCATCTCCCAACCGCAGCGGTAAAGACACCAAGGTCCCGGCGCATACGAC 661
DB 2329 CTCACCTCAGTCTCCCAACGSCAGCGGCAAAAGACCAAGGTCCCGGCTGATATGAG 2388
QY 662 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCCGTTGTCTCAACAATTTGGCTT 721
DB 2389 CTCAGGGCTATAAGGTGTGTGTCTCAACCCCTCTGTTGCTCAACAACCTGGGCTT 2448
QY 722 CTTACATGTCCAAGGCCCATGGATGTATCTTAACATCAGGACTGGGTGAGGACAATTA 781
DB 2449 CTTACATGTCCAAGGCTCATGGATGTATCTTAACATCAGGACTGGGTGAGGACAATTA 2508
QY 782 CTACTGGCAGCCCGATCAGTATTCACCTACGCAAGTTCCTTCCGACGCGCGGTGT 841
DB 2509 CCACCTGGCAGCCCATCAGTATTCACCTACGCAAGTTCCTTCCGACGCGCGGTGT 2568
QY 842 CAGGGGTGTCTATCACAATAATTTGTGACGATGTCCATCTCCACGGATGCAACATCA 901
DB 2569 CCGGGGGGCGCTTATGACATAATAATTTGTGACGATGTCCATCTCCACGGATGCAACATCA 2628
QY 902 TCTTGGGCATTTGGCACTGTCTTACCAAGCAGACGCGGGGGGAGACTGACTGTGC 961
DB 2629 TCTTGGGCATTCGGCATGTCTTACCAAGCAGACGCGGGGGGAGACTGACTGTGC 2688
QY 962 TCGCCACCGCTACCCCTCTCGGCTCGCTACCTGTGCTCCCATCTTAACATCGAGGAGT 1021
DB 2689 TCGCCACCGCCACCCCTCTCGGCTCGCTACCTGTGCTCCCATCTTAACATCGAGGAGT 2748
QY 1022 CTTCTGTCACTACCGGAGATGCCCTTTTATGGCAAGCTATTCCTTTGAAGCAATTA 1081
DB 2749 CTTCTGTCACTACCGGAGATGCCCTTTTATGGCAAGCTATTCCTTTGAAGCAATTA 2808
QY 1082 AGGGGGGAGACATCTCTCTCCGCTCAAGAGAGTGCAGCAGCTCGCGCAA 1141
DB 2809 AGGGGGGAGACATCTCTCTCCGCTCAAGAGAGTGCAGCAGCTCGCGCAA 2868
QY 1142 AACTGTGTGCGGTGGCGCTCAATGCGCTGTCTTACTACCGCGCTTGTGTGCTGCTCA 1201
DB 2869 AGCTGTGCTATGGGCATCANTGCGTGGCTCTACTACCGCGCTTGTGTGCTGCTGCTCA 2928
QY 1202 TCCCGACAGTGGTGAAGTGTGTGCTGGGCAACTGAGCGCTCATGACCGGCTTTACCG 1261
DB 2929 TCCCGACAGTGGTGAAGTGTGTGCTGGGCAACTGAGCGCTCATGACCGGCTTTACCG 2988
QY 1262 GGGACTTCGATTCGGTGTAGTGCACACCTGTGTCAACAGAGTGCAGCTTCAGCC 1321
DB 2989 GGGACTTCGATTCGGTGTAGTGCACACCTGTGTCAACAGAGTGCAGCTTCAGCC 3048
QY 1322 TTGACCTTACCTTACCATTTGAGACAATCACGCTTCCCGAGGATGTGTCTCCGCTACTC 1381
DB 3049 TTGACCTTACCTTACCATTTGAGACAATCACGCTTCCCGAGGATGTGTCTCCGCTACTC 3108
QY 1382 AACGTCGGGGTAGGACTGGCAGAGGAGGAGGAGGATCTACAGATTTGTGGCACCGGGGG 1441
DB 3109 AACGTCGGGGTAGGACTGGCAGAGGAGGAGGAGGATCTACAGATTTGTGGCACCGGGGG 3168
QY 1442 AGCGTCTTCTGCGCATGTTTGTGCTGTCTCTGCGAGTGTATGACGCGGGTGTG 1501
DB 3169 AGCGTCTTCTGCGCATGTTTGTGCTGTCTCTGCGAGTGTATGACGAGGCTGTG 3228
QY 1502 CTTGGTATGAGCTTACGCGCCCGCGAGACCAAGCTTAGGCTACGAGCATATGAAACACCC 1561

Db 3229 CTTGGTATGAGCTCAGCGCCGCGAGACTACGTTAGGCTACAGGCTACATGAACACCC 3288
QY 1562 CGGACTTCCGCTGTCAGAGACATCTTGAATTTTGGAGGCGCTTTTACGGGCTCA 1621
Db 3289 CGGGCTTCCGCTGTCAGAGACATCTTGAATTTTGGAGGCGCTTTTACAGGCTCA 3348
QY 1622 CCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
Db 3349 CTCATATAGATGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTACC 3408
QY 1682 TGTAGGTACCAAGCCAGCTGTGCTAGAGCTCAAGCCCTCCCGCTGTGGAGCC 1741
Db 3409 TGTAGGTACCAAGCCAGCTGTGCTAGAGCTCAAGCCCTCCCGCTGTGGAGCC 3468
QY 1742 AGATGTGAAGTCTTGATCCGCTCAAGCCACCTCCATGGGCCCAACACCTCTGCTAT 1801
Db 3469 AGATGTGAAGTCTTGATCCGCTCAAGCCACCTCCATGGGCCCAACACCTCTGCTAT 3528
QY 1802 ATAGACTGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAAATATCA 1861
Db 3529 ACAGACTGGCGCTGTTCAGAAATGAAGTCAACCTGACGACCCAGTCAACAAATATCA 3588
QY 1862 TGACATGTATCGGCTGACCTGAGGTGCTACAGATACCTGGGTGCTGCTGGCGGCG 1921
Db 3589 TGACATGCATGTCGGCGGACCTGGAGGTGCTACAGAGCACTGGGTGCTGCTGGCGGCG 3648
QY 1922 TTCTGCTGCTTGGCGGCTATGCTATCCACAGCTGCGGTGCTAGTAGTAGGA 1981
Db 3649 TCTGCTGCTTGGCGGCTATGCTATCCACAGCTGCGGTGCTAGTAGTAGGA 3708
QY 1982 TTCTGCTGCTGCGGAAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTCG 2041
Db 3709 TCTGCTGCTGCGGAAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGGAGTTCG 3768
QY 2042 ATGAATGAAGAGTGTCT 2059
Db 3769 ATGAGATGAAGAGTGTCT 3786

RESULT 8

AAQ98221 standard; cDNA to mRNA; 7310 BP.

XX AAQ98221;

AC 25-MAR-2003 (revised)

DT 15-AUG-1996 (first entry)

XX AAQ98221;

DE Hepatitis C virus clone genome.

XX Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;

KW antibodies; immunoprophylaxis; sera; serum; ds.

XX Hepatitis C virus.

OS US5443965-A.

XX 22-AUG-1995.

PN 05-APR-1991; 91US-00681703.

XX 06-APR-1990; 90US-00505611.

PR 09-OCT-1990; 90US-00594854.

XX (GENE-) GENELABS INC.

PA Moeckli R, Reyes GR, Kim JP;

XX WPI; 1995-302120/39.

DR New nucleic acids encoding hepatitis C virus antigens - used to develop

PT prods. for detection of HCV-infected sera and prodn. of vaccines and anti

PT -HCV antibodies.

XX Example 4; Fig 11; 71pp; English.

XX Hepatitis C virus (HCV) antigens can be used for detecting HCV infected
CC sera and individuals infected with HCV. They can also be used in an anti-
CC HCV vaccine or for the production of anti-HCV antibodies which can be
CC used for passive immunoprophylaxis. The antigens consistently identify
CC more HCV positive serum samples with a high degree of specificity. See
CC AAQ98202-14 and AAR81939-51. (Updated on 25-MAR-2003 to correct Pf
CC field.) (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 7310 BP; 1494 A; 2217 C; 2060 G; 1539 T; 0 U; 0 Other;

QY Query Match 86.7%; Score 1786; DB 2; Length 7310;

Best Local Similarity 91.7%; Pred No. 0;

Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCCTATCACGGCCTATGCCAGCAGACAAAGGGCCTTTTGGGATGCATAATCACCA 61

Db 1729 TGGCGCCTATCACGGCCTATGCCAGCAGACAAAGGGCCTTTTGGGATGCATAATCACCA 1788

QY 62 GCTTGAACCGCCGCGACAAAACACAGGTGGAGGTGAGTTTCAGATCGTGTCACTGCTG 121

Db 1789 GCTTGAACCGCCGCGACAAAACACAGGTGGAGGTGAGTTTCAGATCGTGTCACTGCTG 1848

QY 122 CCCAGACTTTCTTTGGCAACCTGTCATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCGG 181

Db 1849 CCCAGACTTTCTTTGGCAACCTGTCATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCGG 1908

QY 182 GAACAAAGGACCAATTCGCTCACCTAAGGGTCTCTGTTATCCAGATGTACACCAATGTGGACC 241

Db 1909 GAACAAAGGACCAATTCGCTCACCTAAGGGTCTCTGTTATCCAGATGTACACCAATGTGGACC 1968

QY 242 AAGACTCTGTAGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGGCG 301

Db 1969 AAGACTCTGTAGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGGCG 2028

QY 302 GCTCTCTCGGACCTTTTACCTGCTCAGAGGACGCGCGATGTCACTTCTGTGCGCCGACGGG 361

Db 2029 GCTCTCTCGGACCTTTTACCTGCTCAGAGGACGCGCGATGTCACTTCTGTGCGCCGACGGG 2088

QY 362 GTGATGGCAGGGCGACCTGCTTTTCGCCCGCGCCTATCTCTTACTTTGAAAGCTCTCTCGG 421

Db 2089 GTGATGGCAGGGCGACCTGCTTTTCGCCCGCGCCTATCTCTTACTTTGAAAGCTCTCTCGG 2148

QY 422 GAGGCCCTCTGCTGTGCGCCGCGAGACATGCGGTAGGCATATTCAGAGCCGCGGTATGCA 481

Db 2149 GAGGCCCTCTGCTGTGCGCCGCGAGACATGCGGTAGGCATATTCAGAGCCGCGGTATGCA 2208

QY 482 CCGCTGGAGTGGCTAAGGGGCTGGGACTTTCATCCCGTAGAGCTTAGAGCAACCATGA 541

Db 2209 CCGCTGGAGTGGCTAAGGGGCTGGGACTTTCATCCCGTAGAGCTTAGAGCAACCATGA 2268

QY 542 GGTCCCGCGGTGTTCTCAGACAACTCCTCCCAACAGCAGTGCCTCCAGAGCTTACCAAGTGG 601

Db 2269 GGTCCCGCGGTGTTCTCAGACAACTCCTCCCAACAGCAGTGCCTCCAGAGCTTACCAAGTGG 2328

QY 602 CCCACTGCATGCTCCCAACCGGACGCTAAGACCAAGGTCCCGCGCGCATACGAG 661

Db 2329 CTCACCTGCATGCTCCCAACCGGACGCTAAGACCAAGGTCCCGCGCGCATACGAG 2388

QY 662 CTCAGGGCTACAGGTGCTGCTCAACCCCTCCGTTGCTGCAACATGGCTTTGGTG 721

Db 2389 CTCAGGGCTACAGGTGCTGCTCAACCCCTCCGTTGCTGCAACATGGCTTTGGTG 2448

QY 722 CTTACATGTCCAAGGCCCATGGATTTGATCTTAACATCAGGACTGGGTGAGGACAAATTA 781

Db 2449 CTTACATGTCCAAGGCCCATGGATTTGATCTTAACATCAGGACTGGGTGAGGACAAATTA 2508

QY 782 CTACTGGAGCCCGCATCATCGTATTCACCTACGGCAAGTTCCTTCCGACGCGGGTGT 841

Db 2509 CCACCTGGAGCCCGCATCATCGTATTCACCTACGGCAAGTTCCTTCCGACGCGGGTGT 2568

QY 842 CAGGGGTCCTTATGACATAAATTTGTGACAGTGGCCACTCCACGGATGCAATCCA 901
DB 2569 CGGGGGGCGCTTATGACATAAATTTGTGACAGTGGCCACTCCACGGATGCAATCCA 2628
QY 902 TCTTGGGCATGCGACTGTCCTTGTGACAAAGCAGAGACCGCGGGGCGAGACTGTGTC 961
DB 2629 TCTTGGGCATGCGACTGTCCTTGTGACAAAGCAGAGACTCGGGGGCGAGACTGTGTC 2688
QY 962 TCGCCACCGCTACCCCTCCGGGTCCTGTCACATGTGCCCCCATCTTAACATCGAGAGGTTG 1021
DB 2689 TCGCCACCGCCACCCCTCCGGGTCCTGTCACATGTGCCCCCATCTTAACATCGAGAGGTTG 2748
QY 1022 CTCTGTCCACTACCGGAGATCCCTTTTATGGCAAGCTATTCCTTTGAAGCAATTA 1081
DB 2749 CTCTGTCCACTACCGGAGATCCCTTTTATGGCAAGCTATTCCTTTGAAGCAATTA 2808
QY 1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAAGAAAGTGGCAGAGCTCGCGCAA 1141
DB 2809 AGGGGGGAGACATCTCATCTTCTGCACTCAAAGAAAGTGGCAGAGCTCGCGCAA 2868
QY 1142 AACTGTGCGTGTGGCGCTCAATGCGGTGCTTATACCGGGCTTGTGATGTGCTCA 1201
DB 2869 AGTGTGCTGATTTGGGCATCAATGCGGTGCTTATACCGGGCTTGTGATGTGCTCA 2928
QY 1202 TCCCGACCACTGCTGAGTGTGCTGCTGCACTGACCGCTCATGACCGGCTTTACCG 1261
DB 2929 TCCCGACCACTGCTGAGTGTGCTGCTGCACTGACCGCTCATGACCGGCTTTACCG 2988
QY 1262 GCGACTTCGATTCGGGTAGACTGCAACACGCTGTGTACCCAGACAGCTCGACTTCAGCC 1321
DB 2989 GCGACTTCGATTCGGGTAGACTGCAACACGCTGTGTACCCAGACAGCTCGACTTCAGCC 3048
QY 1322 TTGACCCCTACCTTACCAATGAGACAAATCAAGCTTCCCGAGTGTGTCTCCGCTACTC 1381
DB 3049 TTGACCCCTACCTTACCAATGAGACAAATCAAGCTTCCCGAGTGTGTCTCCGCTACTC 3108
QY 1382 AAGTCGGGGTAGGACTGCGAGGGAAGCCAGGCACTACAGATTTGGCAGCGGGG 1441
DB 3109 AAGTCGGGGTAGGACTGCGAGGGAAGCCAGGCACTACAGATTTGGCAGCGGGG 3168
QY 1442 AGCGTCTCTTGTGCACTGTTGACTCTGCTCTCTGCGAGTGTGTATGACGGGGTTGTG 1501
DB 3169 AGCGCCCTCCGCACTGTTGCACTCTGCTCTCTGCGAGTGTGTATGACGAGGCTGTG 3228
QY 1502 CTTGTGATGACTTACCGCCGCGAGACCAAGTGTAGGCTACGAGTATACATGACACC 1561
DB 3229 CTTGTGATGACTTACCGCCGCGAGACTACAGTGTAGGCTACGAGGCTACATGACACC 3288
QY 1562 CGGACTCTCCGCTGCGCAAGACCATCTTGAATTTGGGAGGCGCTTTTACGGGTCTCA 1621
DB 3289 CGGGGCTTCCGCTGCGCAGGACCATCTTGAATTTGGGAGGCGCTTTTACGGGCTCA 3348
QY 1622 CCCACATAGACGCCCACTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
DB 3349 CTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC 3408
QY 1682 TGTAGGATACCAAGCCACCGTGTGCTGCTAGACTCAAGCCCTCCCGCTGTGGGACC 1741
DB 3409 TGTAGGATACCAAGCCACCGTGTGCTGCTAGAGCTCAAGCCCTCCCGCTGTGGGACC 3468
QY 1742 AGATGTGAAGTCTGTGATCCGCTCAAGCCCACTCTCATGGGCGCAACACCTCTCTAT 1801
DB 3469 AGATGTGAAGTCTGTGATCCGCTCAAGCCCACTCTCATGGGCGCAACACCTCTCTAT 3528
QY 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCACTTACGACCCAGTCAACAAATATCA 1861
DB 3529 ACAGACTGGGCGCTGTTCAGAAATGAAGTCACTTACGACCCAGTCAACAAATATCA 3588
QY 1862 TGACATGTATGCGGCTGACCTGGAGTGTGTACGAGTACCTGGGTCTCTGTGGGGGG 1921
DB 3589 TGACATGTATGCGGCGGACCTGGAGTGTGTACGAGTACCTGGGTCTCTGTGGGGGG 3648

QY 1922 TTCTGGCTGCTTTGGCGCGCTATTGCTTCATCCACAGGCTGCGTGGTTCATAGTAGTAGGA 1981
DB 3649 TCCTGGCTGCTTTGGCGCGCTATTGCTTCATCCACAGGCTGCGTGGTTCATAGTAGTAGGA 3708
QY 1982 TTGTCTTCTCGGAAAGCCGCAATCATACCGACAGGAAAGTCTCTTACCGGAGTTTCG 2041
DB 3709 TCGTCTTCTCGGAAAGCCGCAATCATACCTGACAGGAAAGTCTCTTACCGAGAGTTTCG 3768
QY 2042 ATGAATATGAAGAGTGTCT 2059
DB 3769 ATGAGATGAAGAGTGTCT 3786

RESULT 9

AAA75296

ID AAA75296 standard; cDNA; 8316 BP.

XX

AC AAA75296;

XX 15-JAN-2001 (first entry)

XX

XX cDNA sequence compiled Hepatitis C virus cDNA clones.

DE Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;

KW viral infectivity; viral replication; ds.

XX Hepatitis C virus.

XX

FH Key Location/Qualifiers

FT CDS 1..8316

FT /*tag= a

FT /note= "partial sequence; no termination codon given"

XX

PN EP1034785-A2.

XX

PD 13-SEP-2000.

XX

PF 16-MAR-1990; 2000EP-00109602.

XX

PR 17-MAR-1989; 89US-00325338.

PR 20-APR-1989; 89US-00341334.

PR 18-MAY-1989; 89US-00355002.

PR 16-MAR-1990; 90EP-00302866.

XX (CHIR) CHIRON CORP.

XX

XX Houghton M, Choo Q, Kuo G;

XX

XX WPI; 2000-566891/53.

XX P-PSDB; AAB18540.

XX

XX Novel composition comprising a hepatitis C virus antisense polynucleotide

XX which is complementary to or corresponds to a sense strand of the virus

XX genome, and selectively hybridizes to it.

XX

XX Example; Fig 16; 75pp; English.

XX

XX The specification describes a pharmaceutical composition which comprises

XX a hepatitis C virus (HCV) antisense polynucleotide. The HCV is

XX characterized by a positive stranded RNA genome which has 40% homology at

XX the polypeptide level to a HCV polyprotein. The antisense polynucleotide

XX binds to cellular polynucleotides which enhance and/or are required for

XX viral infectivity, replicative ability or chronicity. The antisense

XX polynucleotides may also be designed to bind with high specificity, to be

XX of increased stability, to be stable and to have low toxicity. The

XX composition also comprises an agent which causes viral RNA to be

XX inactive. The composition is used for preventing HCV replication in a

XX system. The present sequence represents a novel HCV cDNA sequence, which

XX is used in the course of the invention

XX

SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;

XX

Query Match 86.7%; Score 1786; DB 3; Length 8316;

QY 482 CCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 541
DB 3876 CCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 3935
QY 542 GGTCCCGGTGTTCTCAGACAACCTCCCGACAGAGTGGCCCGAGAGCTACCAAGTGG 601
DB 3936 GGTCCCGGTGTTCTCAGAGATACTCCTCTCCACAGTAGTGGCCCGAGAGCTTCAGAGTGG 3995
QY 602 CCCACCTGCATGCTCCCGACCGGTAAGAGCACCAAGGTCCCGCGCGCATACGCAG 661
DB 3996 CTCACCTCCATGCTCCCGACCGGTAAGAGCACCAAGGTCCCGCGCGCATATGCAG 4055
QY 662 CTCAGGGCTACAGGTGCTGGTCTCAACCCCTCCGTTCGTCACAACTAGGGCTTTGGTG 721
DB 4056 CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTG 4115
QY 722 CTTACATGTCGAAGGCCATGGGATTCCTTAACATCAGAGCTGGGGTGAGACAATTA 781
DB 4116 CTTACATGTCGAAGGCTCATGGGATCGATCCTTAACATCAGAGCTGGGGTGAGACAATTA 4175
QY 782 TCTAGGAGCGGATCAGTATTCACCTACGGCAAGTTCCTTGGCGAGCGCGGTGT 841
DB 4176 CCACTGGAGCGCCATCAGTACTCCACCTACGGCAAGTTCCTTGGCGAGCGCGGTGT 4235
QY 842 CAGGGGTGCTTATGACATAATAATTTGTGACAGTGCACCTCCACGATGCAACATCCA 901
DB 4236 CGGGGGCGCTTATGACATAATAATTTGTGACAGTGCACCTCCACGATGCAACATCCA 4295
QY 902 TCTTGGGCAATGGCACTGCTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 961
DB 4296 TCTTGGGCAATCGGCACTGCTTGACCAAGCAGAGACTCGGGGGGAGACTGGTTGTGC 4355
QY 962 TGGCCACCGTACCCCTCGGGCTCCGTCACTGTGCCCATCTCAACATCGAGGAGTTG 1021
DB 4356 TGGCCACCGTACCCCTCGGGCTCCGTCACTGTGCCCATCTCAACATCGAGGAGTTG 4415
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAAGCAATTA 1081
DB 4416 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAAGTAATCA 4475
QY 1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGGCGAGAGCTCGCGGAA 1141
DB 4476 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGGCGAGAGTGGCGGAA 4535
QY 1142 AACTGTGTCGCTTGGCGCTCAATGCGTGTCTACTACCGCGCTTGTATGTTCGCTCA 1201
DB 4536 AGCTGTGTCGCTTGGCGCTCAATGCGTGTCTACTACCGCGCTTGTATGTTCGCTCA 4595
QY 1202 TCCGACCAAGTGGTGTGCTGTGCTGGCAACTGACCGCCCTCATGACCGGCTTTACCG 1261
DB 4596 TCCGACCAAGTGGTGTGCTGTGCTGGCAACTGACCGCCCTCATGACCGGCTTTACCG 4655
QY 1262 GCGACTTCGATTCGCTGTAGACTGCAACGCTGTGTCAACCGAGAGTGGCTTCAGCC 1321
DB 4656 GCGACTTCGATTCGCTGTAGACTGCAACGCTGTGTCAACCGAGAGTGGCTTCAGCC 4715
QY 1322 TTGACCTTACCTTACCATTTAGACAATCAAGCTTCCCGAGATGTGTCTCCGCTACTC 1381
DB 4716 TTGACCTTACCTTACCATTTAGACAATCAAGCTTCCCGAGATGTGTCTCCGCTACTC 4775
QY 1382 AACGTCGGGGTAGGACTGGCAGAGGAAGCCAGGCAATCTACAGATTTTGGCACCGGGG 1441
DB 4776 AACGTCGGGGTAGGACTGGCAGAGGAAGCCAGGCAATCTACAGATTTTGGCACCGGGG 4835
QY 1442 AGGCTCTTCTGGCATGTTGACTGCTGTCTCTGCGAGTGTCTATGACGGGGTGTG 1501
DB 4836 AGCGCCCTTCCGCAATGTTGCACTGCTGCTCTCTGCTGAGTGTCTATGACGAGGCTGTG 4895
QY 1502 CTTGGTATGAGCTTACCGCGCGGAGACCAAGTGTAGGCTACGAGCATACATGAACACC 1561
DB 4896 CTTGGTATGAGCTTACCGCGCGGAGACCAAGTGTAGGCTACGAGGCTACATGAACACC 4955
QY 1562 CGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACGGGTCTCA 1621

DB 4956 CGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGTCTTTACAGGCTCA 5015
QY 1622 CCCACATAGACGCCCACTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
DB 5016 CTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCCCTACC 5075
QY 1682 TGGTAGCGTACCAAGCCACCGGTGCTGCTAGAGCTCAAGCCCTCCCGTGTGGGACC 1741
DB 5076 TGGTAGCGTACCAAGCCACCGGTGCTGCTAGAGCTCAAGCCCTCCCGTGTGGGACC 5135
QY 1742 AGATGTGAAGTGTGATTCGCTCAAGCCCAACCTTCATGGGCAACACCTCTGCTAT 1801
DB 5136 AGATGTGAAGTGTGATTCGCTCAAGCCCAACCTTCATGGGCAACACCTCTGCTAT 5195
QY 1802 ATAGACTGGCGCTGCTCAGATGAAGTCAACCTCAGCACCACCTCACCAGTATATCA 1861
DB 5196 ACAGACTGGCGCTGCTCAGATGAAGTCAACCTCAGCACCACCTCACCAGTATATCA 5255
QY 1862 TGACATGATGTGGCTGCTACCTGGAGGTGCTCAAGTACCTGGGTGCTCTGGCGGCG 1921
DB 5256 TGACATGATGTGGCGGCTGCTGGAGGTGCTCAAGTACCTGGGTGCTCTGGCGGCG 5315
QY 1922 TTTCTGGCTGCTTGGCGGCTGCTGCTATCCATCAGGCTGCTGCTATAGTAGTAGGA 1981
DB 5316 TCTTGGCTGCTTGGCGGCTGCTGCTATCCATCAGGCTGCTGCTATAGTAGTAGGA 5375
QY 1982 TTGTCTTGTCCGGAAGCGGCAATCATCCGACAGGAGTCTCTACCGGAGTTCCG 2041
DB 5376 TGTCTTGTCCGGAAGCGGCAATCATCCGACAGGAGTCTCTACCGGAGTTCCG 5435
QY 2042 ATGAATGAAGAGTGTCT 2059
DB 5436 ATGAGATGAAGAGTGTCT 5453

RESULT 12
AAA75297

ID AAA75297 standard; cdna; 9185 BP.

XX AAA75297;

XX 15-JAN-2001 (first entry)

XX Sense strand of HCV encoding a polyprotein.

XX Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;

XX viral infectivity; viral replication; ds.

XX Hepatitis C virus.

XX Key Location/Qualifiers

XX CDS 320..9184

XX /*tag= a

XX /*note= "partial sequence; no termination codon given"

XX EP1034785-A2.

XX 13-SEP-2000.

XX 16-MAR-1990; 2000EP-00109602.

XX 17-MAR-1989; 89US-00325338.

XX 20-APR-1989; 89US-00341334.

XX 18-MAY-1989; 89US-00355002.

XX 16-MAR-1990; 90EP-00302866.

XX (CHIR) CHIRON CORP.

XX Houghton M, Choo Q, Kuo G;

XX WPI; 2000-566891/53.

XX P-PSDB; AAB18541.

XX Novel composition comprising a hepatitis C virus antisense polynucleotide
PT which is complementary to or corresponds to a sense strand of the virus
PT genome, and selectively hybridizes to it.
XX Example; Fig 17; 75pp; English.

XX The specification describes a pharmaceutical composition which comprises
CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
CC characterized by a positive stranded RNA genome which has 40% homology at
CC the polypeptide level to a HCV polypeptide. The antisense polynucleotide
CC binds to cellular polynucleotides which enhance and/or are required for
CC viral infectivity, replicative ability or chronicity. The antisense
CC polynucleotides may also be designed to bind with high specificity, to be
CC of increased stability, to be stable and to have low toxicity. The
CC composition also comprises an agent which causes viral RNA to be
CC inactive. The composition is used for preventing HCV replication in a
CC system. The present sequence represents a novel HCV cDNA sequence, which
CC is used in the course of the invention

XX Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Query Match 86.7%; Score 1786; DB 3; Length 9185;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY	2	TGGCGCCCTATCAGCGCCCTATGCCAGCAGACAAAGGGGCCCTTTGGGATGCATATCACCA	61
DB	3396	TGAGGCCCATCAGCGCGTAGCCAGCAGACAAAGGGGCCCTTAGGGTGCATATCACCA	3455
QY	62	GCTTGACCGCCCGGACAAAAACAGGTGAGGGTGAAGTTTCAGATCGTCAACTGCTG	121
DB	3456	GCCTAACTGCGCGGACAAAAACAAAGTGAGGGTGAAGTCCAGATTTGTCAACTGCTG	3515
QY	122	CCGAGCTTTCTGGCAACTGCATTAACGGGGTGTGTGGAGTCTACCATGGAGCCG	181
DB	3516	CCCAACCTTCTGGCAACTGCATTAACGGGGTGTGTGGAGTCTACCATGGAGCCG	3575
QY	182	GAACAGGACCATTTGGTGTACCTTAAGGGTCTCTGTATCCAGATGTACCAATGTGGACC	241
DB	3576	GAACAGGACCATTTGGTGTACCTTAAGGGTCTCTGTATCCAGATGTACCAATGTGGACC	3635
QY	242	AAGACCTCTGAGTGGCGCGCTCCCAAGGTGCCCGCTCATTAACCAATGCATTTGG	301
DB	3636	AAGACCTCTGAGTGGCGCGCTCCCAAGGTAGCGCTCATTAACCAATGCATTTGG	3695
QY	302	GCTTCTCGGACCTTTACCTGGTCAAGGACCAAGCGATGTCTCTGTGCGCGGCGG	361
DB	3696	GCTTCTCGGACCTTTACCTGGTCAAGGACCAAGCGATGTCTCTGTGCGCGGCGG	3755
QY	362	GTGATGGCAGGGCAGCTCTTTTCGCCCGGCTATCTCTTACTTGAAGGCTCCTCGG	421
DB	3756	GTGATGGCAGGGCAGCTCTTTTCGCCCGGCTATCTCTTACTTGAAGGCTCCTCGG	3815
QY	422	GAGGCCCTCTGTGTCGCCCGCAGGACATGCCGTAGGCATATTTCAGAGCGCGGTATGCA	481
DB	3816	GAGGCCCTCTGTGTCGCCCGCAGGACATGCCGTAGGCATATTTCAGAGCGCGGTATGCA	3875
QY	482	CCCGTGGAGTGGCTAAGCGGGTGCATTTATCCCGGTAGAGCTTAGAGACAACATGA	541
DB	3876	CCCGTGGAGTGGCTAAGCGGGTGCATTTATCCCGGTAGAGACCTTAGAGACAACATGA	3935
QY	542	GGTCCCGGTTCTCAGACAACTCTTCCCGCAGCAGTGGCCCGCAGAGCTACCAAGTGG	601
DB	3936	GGTCCCGGTTCTCAGACAACTCTTCCCGCAGCAGTGGCCCGCAGAGCTACCAAGTGG	3995
QY	602	CCCACTGCAATGCTCCCGCAGCGGTGAAGAGCACCAAGGTCCCGGCGCATACGCAG	661
DB	3996	CTCACTCCATGCTCCCGCAGCGGTGAAGAGCACCAAGGTCCCGGCGCATATGCAG	4055
QY	662	CTCAGGGCTACAAGGTGCTGGTCTCAACCCCTCCGTTGCTGCAACATGGGCTTTGGTG	721
DB	4056	CTCAGGGCTACAAGGTGCTGGTCTCAACCCCTCTGTTGCTGCAACATGGGCTTTGGTG	4115

QY	722	CTTACATGTCCAGGCCCATGGGATGTGATCTTAACATCAGGACTGGGGTGAGACAATTA	781
DB	4116	CTTACATGTCCAGGCCCATGGGATGTGATCTTAACATCAGGACTGGGGTGAGACAATTA	4175
QY	782	CTACTGGCAGCCGATCAGGTATTCACCTACGGCAAGTTCCTTGGCGAGCGGGTGT	841
DB	4176	CCACTTGGCAGCCGATCAGGTATTCACCTACGGCAAGTTCCTTGGCGAGCGGGTGT	4235
QY	842	CAGGGGGTGTCTATGACATAAATTTGTGACAGGTGCCACTCCACGGATGCAACATCCA	901
DB	4236	CGGGGGGCGCTTATGACATAAATTTGTGACAGGTGCCACTCCACGGATGCAACATCCA	4295
QY	902	TCTTGGGCATTTGCGCACTGTCTTACCAAGCAGAGACCGCGGGGGGAGACTGTGTC	961
DB	4296	TCTTGGGCATTTGCGCACTGTCTTACCAAGCAGAGACTGCGGGGGGAGACTGTGTC	4355
QY	962	TGCGCACCCGTACCCCTCGGGGTCCGTCACTGTGCCCCCATCTTAACATCGAGGAGTTG	1021
DB	4356	TGCGCACCCGTACCCCTCGGGGTCCGTCACTGTGCCCCCATCTTAACATCGAGGAGTTG	4415
QY	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCCTTTGAAGCAATTA	1081
DB	4416	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCCTTTGAAGCAATTA	4475
QY	1082	AGGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAGAAAGTGCAGAGCTCGCGCAA	1141
DB	4476	AGGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAGAAAGTGCAGAGAACTCGCGCAA	4535
QY	1142	AACTGTGTCGTTGGGCGTCAATGCGGTCTTACTACCGGGCTTTGATGTGTCGTCA	1201
DB	4536	AGCTGTGTCGTTGGGCGTCAATGCGGTCTTACTACCGGGCTTTGATGTGTCGTCA	4595
QY	1202	TCCGACACAGTGTGAGCTGTGCTGCGCAACTGACGCCCTCATGACCGGCTTTACCG	1261
DB	4596	TCCGACACAGTGTGAGCTGTGCTGCGCAACTGACGCCCTCATGACCGGCTTTACCG	4655
QY	1262	GCGACTTTCATTCGCTGTAGTGTGCAACACGCTGTGTACCCAGACAGTGCAGCTTCAGCC	1321
DB	4656	GCGACTTTCATTCGCTGTAGTGTGCAACACGCTGTGTACCCAGACAGTGCAGCTTCAGCC	4715
QY	1322	TTGACCTTACCTTACCATTTAGACAAATCAACGTTTCCCGAGATGTGTCTCCGTAATC	1381
DB	4716	TTGACCTTACCTTACCATTTAGACAAATCAACGTTTCCCGAGATGTGTCTCCGTAATC	4775
QY	1382	AACTGTGGGGTAGGACTGGCAGAGGAGCAGGACATCTACAGATTTGGGACCGGGG	1441
DB	4776	AACTGTGGGGTAGGACTGGCAGAGGAGCAGGACATCTACAGATTTGGGACCGGGG	4835
QY	1442	AGCGTCTTCTGCGCATTTGATCTGCTGTCTCTGCGAGTGTATGACCGGGTGTG	1501
DB	4836	AGCGTCTTCTGCGCATTTGATCTGCTGTCTCTGCGAGTGTATGACCGGGTGTG	4895
QY	1502	CTTGTGTAGCTTACGCGCGCGAGACCAACAGATTTAGGCTTACAGACATACAGAACCC	1561
DB	4896	CTTGTGTAGCTTACGCGCGCGAGACTTACAGTTTAGGCTTACAGACATACAGAACCC	4955
QY	1562	CGGGACTTCCCGTGTGCCAAGACCATTTGAATTTTGGAGGGCGTCTTTACGGGTCTCA	1621
DB	4956	CGGGACTTCCCGTGTGCCAAGACCATTTGAATTTTGGAGGGCGTCTTTACGGGTCTCA	5015
QY	1622	CCACATAGACGCCCACTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC	1681
DB	5016	CTCATATAGTGCCTCATCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC	5075
QY	1682	TGCTAGCTTACCAAGCCACCGTGTGCTGTAGAGCTCAAGCCCTCCCGGTGTGGAGC	1741
DB	5076	TGCTAGCTTACCAAGCCACCGTGTGCTGTAGAGCTCAAGCCCTCCCGGTGTGGAGC	5135
QY	1742	AGATGTGGAAGTGTGTGCTCTCAAGCCCACTCTCAAGGGGCAACACCTCTCTCTAT	1801
DB	5136	AGATGTGGAAGTGTGTGCTCTCAAGCCCACTCTCAAGGGGCAACACCTCTCTCTAT	5195

QY 1802 ATAGACTGGGCGTGTCCAGAAATGAAGTCACCTGACGCCACCCAGTCACCAAGTATATCA 1861
DB 5196 ACAGACTGGGCGTGTTCAGAAATGAATCACCCTGACGCCACCCAGTCACCAATATATCA 5255
QY 1862 TGACATGTATGTCCGCTGACCTGGAGGTCGTCAACGAGTACCTGGGTGCTCGTTGGCGGG 1921
DB 5256 TGACATGTATGTCCGCTGACCTGGAGGTCGTCAACGAGTACCTGGGTGCTCGTTGGCGGG 5315
QY 1922 TTCTGGCTCTTTGGCGCGGTATGCTTATCCACAGGTCGCTCATAGTAGTAGGA 1981
DB 5316 TCCTGGCTCTTTGGCGCGGTATGCTTATCCACAGGTCGCTCATAGTAGTAGGA 5375
QY 1982 TTGCTTTGTCGGAAGCGGCAATCATACCCGACAGGGAAGTCCTCTACCGGAGTTGG 2041
DB 5376 TGTCTTTGTCGGAAGCGGCAATCATACCTTACCTTGACAGGGAAGTCCTCTACCGAGATTGG 5435
QY 2042 ATCAAAATGAAGAGTGCT 2059
DB 5436 ATGAGATGAAGAGTGCT 5453

RESULT 13

ADN35979

ID ADN35979 standard; cDNA; 9185 BP.

AC ADN35979;

XX 17-JUN-2004 (first entry)

DT HCV cDNA clone #2.

DE Antiviral; Vaccine; hepatitis C virus infection; HCV infection; ss.

KW Hepatitis C virus.

XX BP1394255-A2.

XX 03-MAR-2004.

XX 16-MAR-1990; 2003EP-00016585.

XX 17-MAR-1989; 89US-00325338.

XX 20-APR-1989; 89US-00341334.

XX 18-MAY-1989; 89US-00355002.

XX 16-MAR-1990; 90EP-00302866.

XX (CHIR) CHIRON CORP.

PA Houghton M, ; Choo Q, Kuo G;

XX WPI; 2004-193149/19.

XX P-PSDB; ADN35978.

XX Novel purified hepatitis C virus polypeptide comprising epitope encoded

PT by hepatitis C virus cDNA, useful as vaccine for treating hepatitis C

PT virus.

XX Claim 1; Fig 17; 79pp; English.

XX The present invention relates to hepatitis C virus (HCV) proteins and

CC cDNA sequences. The sequences are useful in immunoassays for detecting

CC antibodies directed against HCV antigen; preparing host cells transformed

CC with a recombinant polynucleotide; screening antiviral agents and

CC determining the effect of antiviral agent in inhibiting viral replication

CC in cell culture system; and developing vaccine for treating HCV

CC infection.

XX Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

SQ Query Match 86.7%; Score 1786; DB 12; Length 9185;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTTATCACGCGCTATGCCAGCAGACAAGGGGCTTTTGGATGCAATATCACCA 61
DB 3396 TGGCGCTTATCACGCGCTATGCCAGCAGACAAGGGGCTTTTGGATGCAATATCACCA 3455
QY 62 GCTTGACCGGCGGGCAAAAACCAAGGTGAGGTGAGGTTCAGATCGTGTCAACTGTG 121
DB 3456 GCCTAACTGGCGGGCAAAAACCAAGGTGAGGTGAGGTTCAGATCGTGTCAACTGTG 3515
QY 122 CCCAGACTTTCTTGGCAACCTTGCAATTAACGGGTGTTGACTGTCTACCATGAGGCG 181
DB 3516 CCCAAACCTTCTTGGCAACCTTGCAATTAACGGGTGTTGACTGTCTACCATGAGGCG 3575
QY 182 GAAACAAGACCAATTCGCTCACCTTAAGGTCCTGTTATCCAGATGTACCAATGTGACC 241
DB 3576 GAAACAAGACCAATTCGCTCACCTTAAGGTCCTGTTATCCAGATGTACCAATGTGACC 3635
QY 242 AAGACTGCTAGGTGGCGCGCTCCCAAGGTGCGGCTCATTAACACATGCACTTGG 301
DB 3636 AAGACTGCTAGGTGGCGCGCTCCCAAGGTGCGGCTCATTAACACATGCACTTGG 3695
QY 302 GCTCCTCGGACCTTTTACCTGCTCAGAGGACGCGGATGTATTCTCTGTGGCGGACGCG 361
DB 3696 GCTCCTCGGACCTTTTACCTGCTCAGAGGACGCGGATGTATTCTCTGTGGCGGACGCG 3755
QY 362 GTGATGGCAGGGGACGCTGCTTTTGGCGCGGCTTATCTCTTACTTTGAAAGGCTCCTCG 421
DB 3756 GTGATGGCAGGGGACGCTGCTTTTGGCGCGGCTTATCTCTTACTTTGAAAGGCTCCTCG 3815
QY 422 GAGGCTCTGCTGTGCGCGGACGACATGCGGTAGGATATTCAGAGCGGGTATGCA 481
DB 3816 GAGGCTCTGCTGTGCGCGGACGACATGCGGTAGGATATTCAGAGCGGGTATGCA 3875
QY 482 CCGGTGAGGTGGCTAAGGCGGTGGACTTATCCCGGTAGAGAGCTTAGAGACAACCATGA 541
DB 3876 CCGGTGAGGTGGCTAAGGCGGTGGACTTATCCCGGTAGAGAGCTTAGAGACAACCATGA 3935
QY 542 GGTCCCCGGTGTCTCAGACAACTCCTCCCAACAGCAGTGCGCCCGAGAGCTTACCAAGTGG 601
DB 3936 GGTCCCCGGTGTCTCAGACAACTCCTCCCAACAGCAGTGCGCCCGAGAGCTTACCAAGTGG 3995
QY 602 CCCACTGCTGCTCCCGGCGGCTAAGAGCACCAGGTCCTCCGCGGACATACGAG 661
DB 3996 CTCACCTCCATGCTCCCGGCGGCTAAGAGCACCAGGTCCTCCGCGGACATACGAG 4055
QY 662 CTCAGGCTTACAAGGTGCTGTGCTCAACCCCTCCGCTGTGTCGCAACAATGGGCTTTGGTG 721
DB 4056 CTCAGGCTTACAAGGTGCTGTGCTCAACCCCTCCGCTGTGTCGCAACAATGGGCTTTGGTG 4115
QY 722 CTTACATGTCCAAGGCCCATGGAATTGATCTTAAACATCAGAGCTGGGTGAGGACAATTA 781
DB 4116 CTTACATGTCCAAGGCCCATGGAATTGATCTTAAACATCAGAGCTGGGTGAGGACAATTA 4175
QY 782 CTACTGGCAGCCCGATCAGTATTCACCTAGCGAAGTTCCTCCGACGCGGGTGT 841
DB 4176 CCACCTGGCAGCCCGATCAGTATTCACCTAGCGAAGTTCCTCCGACGCGGGTGT 4235
QY 842 CAGGGGTGCTTATGACATAATAATTTGTGAGAGTGGCCACTTCCACGGATGCAACATCCA 901
DB 4236 CAGGGGTGCTTATGACATAATAATTTGTGAGAGTGGCCACTTCCACGGATGCAACATCCA 4295
QY 902 TCTTGGGCAATTGGCATGCTGCTTGAACAGAGAGACCGGGGGGAGAGCTGACTGTGC 961
DB 4296 TCTTGGGCAATTGGCATGCTGCTTGAACAGAGAGACTGCGGGGGGAGAGCTGACTGTGC 4355
QY 962 TCGCCACCGCTACCCCTCGGGCTCCGTCACGTGCGCCATCTTAACATCGAGGAGGTG 1021
DB 4356 TCGCCACCGCTACCCCTCGGGCTCCGTCACGTGCGCCATCTTAACATCGAGGAGGTG 4415
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTGAAGCAATTA 1081
DB 4416 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTGAAGCAATTA 4475
QY 1082 AGGGGGGGAGACATCTCATCTCTGCGCACTCAAGAGAGTGGCAGAGCTCGCGCA 1141

```
Db 4476 AGGGGGGAGACATCTCATCTTCTGTCAATCAAGAGAAAGTGCAGCAACTCGCGCAA 4535
Qy 1142 AACTGTGCGGTGGGGCGTCAATCCGCTGTACTACCGCGCCCTTGATGTGTCGGTCA 1201
Db 4536 AGCTGTGCGCATTTGGGCATCAATGCCGTGACCTACTACCGCGCTCTTGAGTGTCCGTCA 4595
Qy 1202 TCCCGACAGTGTGACGTTGTGTCGTGTGGCAATGACGCCCTCATGACCGGCTTACCG 1261
Db 4596 TCCCGACAGCGCATTTGTGTCGTGTGGCAACCGATGCCCTCATGACCGGCTATACCG 4655
Qy 1262 GCGACTTCGATTCGCTGATAGATGCAACACGCTGTGTCACCCAGACAGTGCATTCAGCC 1321
Db 4656 GCGACTTCGACTCGGTGATAGATGCAATACGCTGTGTCACCCAGACAGTGCATTCAGCC 4715
Qy 1322 TTGACCTTACCTTACCATTGAGACAATCACGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381
Db 4716 TTGACCTTACCTTACCATTGAGACAATCACGCTTCCCGAGGATGCTGTCTCCCGCACC 4775
Qy 1382 AACGTGGGGTAGGACTGGCAGAGGAGACCGAGCATCTACAGATTTGTGGCAACCGGGG 1441
Db 4776 AACGTGGGGTAGGACTGGCAGAGGAGACCGAGCATCTACAGATTTGTGGCAACCGGGG 4835
Qy 1442 AGCGTCTCTTGGCATGTTTGTGCTGCTGCTCTGCGAGTGTGTCAGCGGGTGTG 1501
Db 4836 AGCGCTCTTGGCATGTTTGTGCTGCTGCTCTGCGAGTGTGTCAGCGGGTGTG 4895
Qy 1502 CTTGGTATGAGCTTACCGCCCGAGACCAACAGTGTAGGCTACGAGCATATCAACACCC 1561
Db 4896 CTTGGTATGAGCTTACCGCCCGAGACCAACAGTGTAGGCTACGAGCATATCAACACCC 4955
Qy 1562 CGGAGCTTCCCGTGTGCAAGACAATCTTGAATTTTGGAGGCGCTTTACCGGTCTCA 1621
Db 4956 CGGAGCTTCCCGTGTGCAAGACAATCTTGAATTTTGGAGGCGCTTTACAGGCTCA 5015
Qy 1622 CCACATAGAGCGCCACTTCTATCCAGACAAGCAGAGTGGGGAACCTTCCCTATC 1681
Db 5016 CTATATAGATGGCCACTTCTATCCAGACAAGCAGAGTGGGGAACCTTCCCTATC 5075
Qy 1682 TGTAGCTTACCAAGCACCCTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
Db 5076 TGTAGCTTACCAAGCACCCTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 5135
Qy 1742 AGATGGGAAGTCTTGATCCGCTCAAGCCCACTTCCATGGGCCAACACCTTCTCTAT 1801
Db 5136 AGATGGGAAGTCTTGATCCGCTCAAGCCCACTTCCATGGGCCAACACCTTCTCTAT 5195
Qy 1802 ATAGACTGGGCGTGTCCAGAAATGAAGTCACTCCAGCAGCACCAGTCAACCAATATCA 1861
Db 5196 ACAGACTGGGCGTGTTCAGAAATGAATCACTCCAGCAGCACCAGTCAACCAATATCA 5255
Qy 1862 TGACATGTATGTGGGCTGACCTGAGTGTGTACGAGTACTCTGGGTGCTGTTGGCGGG 1921
Db 5256 TGACATGTATGTGGGCGACCTGAGAGTGTGTACGAGCACCCTGGGTGCTGTTGGCGGG 5315
Qy 1922 TTTGCTGCTGTTGGCGCGTATTTGCTATFCCAAGGCTGCGGTGCTATAGTAGGTAGGA 1981
Db 5316 TTTGCTGCTGTTGGCGCGTATTTGCTGCTCAACAGGCTGCGGTGCTATAGTAGGTAGGA 5375
Qy 1982 TTTGCTGCTGTTGGCGCGGCAATCATATCCAGCAGGGAAGTCTCTACCGGGAGTTGG 2041
Db 5376 TTTGCTGCTGTTGGCGGCAATCATATCCAGCAGGGAAGTCTCTACCGGGAGTTGG 5435
Qy 2042 ATGAAATGGAAGAGTGTCT 2059
Db 5436 ATGAGATGGAAGAGTGTCT 5453
```

RESULT 14
AAT12710
ID AAT12710 standard; cDNA; 9401 BP.
XX
AC AAT12710;

```
XX 25-MAR-2003 (revised)
DT 15-MAY-1996 (first entry)
XX Hepatitis C virus polyprotein.
XX Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;
KW antibodies; ds.
XX Hepatitis C virus.
XX Key Location/Qualifiers
FT CDS 342..9378
FT /*tag= a
XX
PN BP693687-A1.
XX
XX 24-JAN-1996.
XX
XX 03-APR-1991; 95BP-00114016.
XX
XX 04-APR-1990; 90US-00504352.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Houghton M, Choo Q, Kuo G;
PI
XX WPI; 1996-117956/13.
XX
XX P-PSDB; AAR90931.
XX
XX Combinations of synthetic Hepatitis C Virus antigens - provide more
XX effective diagnosis of Non-A, Non-B Hepatitis.
XX
XX Disclosure; Fig 1(A-Y); 53pp; English.
XX
XX The combination comprises an HCV antigen from the C domain (pref. C22 -
XX AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -
XX AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5
XX (AAR90934) domain. The antigens may in the form of a fusion protein, a
XX simple physical mixture, or the individual antigens commonly bound to a
XX solid matrix. They are pref. prepd. by recombinant DNA techniques
XX (primers are given in AAT12711-T12716), but can be synthesised or
XX isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003
XX to correct PF field.)
XX
XX Sequence 9401 BP; 1883 A; 2673 G; 1985 T; 0 U; 0 Other;
XX
Query Match 86.7%; Score 1786; DB 2; Length 9401;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
Qy 2 TGGCGGCTATCACGGCTATGCCAGCAGACAGAGGGGCTTTTGGGATGCATATCACCA 61
Db 3418 TGGCGGCTATCACGGCTATGCCAGCAGACAGAGGGGCTCTTAGGGTGCATATCACCA 3477
Qy 62 GCTTGACCGCGCGGACAAAAACAGGTGGAGGTGAGGTTTCAGATCGTCAACTGCTG 121
Db 3478 GCCTAACTGCGCGGACAAAAACAGGTGGAGGTGAGGTTTCAGATCGTCAACTGCTG 3537
Qy 122 CCCAGACTTTCTTGGCAACTGCATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCCG 181
Db 3538 CCCAAACCTTCTTGGCAACTGCATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCCG 3597
Qy 182 GAACAAGGACCATTTGCTCACCTCAAGGGTCTTGTATCCAGATGTACACCAATGTGGACC 241
Db 3598 GAACGAGGACCATTTGCTCACCTCAAGGGTCTTGTATCCAGATGTATACCAATGTAGACC 3657
Qy 242 AAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGGC 301
Db 3658 AAGACCTCGTAGGCTGGCCCGCTCCCAAGGTAGCGCTCATTTGACACCTGCACTTGGC 3717
Qy 302 GCTCTCTGGAACCTTACCTTGTGTCAGAGGACGCCGATGTCTTCTGTGCGCGGACGGG 361
```

Db 3718 GCTCCTCGGACCTTTTACCTGGTACAGAGGACGCGAGTGTATTCCTCCGTCGCGCGGGG 3777
QY 362 GTGATGGGAGGGGAGCCCTGCTTTTGGCCCCGGGCTATCTCTTACTTTGAAGGCTCTCTGG 421
Db 3778 GTGATAGCAGGGGAGCCCTGCTGTGCGCCCCGGCCCATTTTCTTACTTTGAAGGCTCTCTGG 3837
QY 422 GAGGCCCTCTGCTGTGCCCCGAGAGCATGCGTAGGCAATTTACAGAGCGCGGTATGCA 481
Db 3838 GGGGTCCGCTGTGTGCCCCGCGGGGACGCGTGGGCATATTTAGGGCCGCGGTGCA 3897
QY 482 CCCGTGAGTGGCTAAGCGGTGGACTTCCATCCCCGTAGAGCTTTAGAGCAACCATGA 541
Db 3898 CCGGTGAGTGGCTAAGCGGTGGACTTTATCCCTGTGAGAACTTAGAGCAACCATGA 3957
QY 542 GGTCCCCGGTGTCTCAGACAATCTCTCCCAACAGCAGTGTGCCCCAGAGCTTCCAGTGG 601
Db 3958 GGTCCCCGGTGTCTCAGGATACTCTCTCCCAAGTAGTGTGCCCCAGAGCTTCCAGTGG 4017
QY 602 CCACCTGATGCTCCACCGGAGCGGTAAAGAGCAACCAAGGTCCCGGCGCATACGCAG 661
Db 4018 CTCACCTCCATGCTCCACAGGAGCGGCAAGAGCAACCAAGGTCCCGCTGCATATGCG 4077
QY 662 CTCAGGGCTACAGGTGCTGGTCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 721
Db 4078 CTCAGGGCTAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACCTGGGCTTTGGTG 4137
QY 722 CTTCATGTCCAAAGGCCATGGATGTGATCTTAACATCAGGACTGGGTGAGGCAATTA 781
Db 4138 CTTCATGTCCAAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAATTA 4197
QY 782 CTACTGGGACCGCATCAGTATTCACCTAGCGCAAGTTCCTTGGCCGAGCGGGTGT 841
Db 4198 CCACTGGGACCGCCATCAGTACTCACTCACTA CCGCAAGTTCCTTGGCGAGCGGGTGT 4257
QY 842 CAGGGGTCTTATGACATAATAATTTGTGACAGTGCACCTCCAGAGTGAACATCA 901
Db 4258 CGGGGGCGCTTATGACATAATAATTTGTGACAGTGCACCTCCAGGATGCCACATCA 4317
QY 902 TCTTGGGCAATGGCACTGCTTACCAAGCAGAGACCGGGGGGAGACTGACTGTGC 961
Db 4318 TCTTGGGCAATGGCACTGCTTACCAAGCAGAGACTCGGGGGGAGACTGTTGTGC 4377
QY 962 TCGCCACCGCTACCCCTCCGGGCTCCGTCACCTGTCCTCCATCTTAACATCGAGAGGTTG 1021
Db 4378 TCGCCACCGCCACCCCTCCGGGCTCCGTCACCTGTCCTCCATCCCAACATCGAGAGTTG 4437
QY 1022 CTCTGTCCACTACCGGAGATCCCTTTTATGGCAAGGTATTCCTTGAAGCAATTA 1081
Db 4438 CTCTGTCCACACCGGAGATCCCTTTTACCGCAAGGTATTCCTTGAAGCAATTA 4497
QY 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGTGCAGAGCTCGCGCAA 1141
Db 4498 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGTGCAGAGCTCGCGCAA 4557
QY 1142 AACTGTTCGGTTCGGGCTCAATTCGGTGTCTTACTACCGGGCTTGTGTCCGTCA 1201
Db 4558 AGTGTTCGATTCGGGCAATTCGGTGTCTTACTACCGGGCTTGTGTCCGTCA 4617
QY 1202 TCCCGACCAAGTGTGAGTGTGCTGTGGCAATCAAGCGCTCATGACGGCTTTACCG 1261
Db 4618 TCCCGACCAAGGCGATGTTGTGCTGGCAACCGATGCCCTCATGACGGCTTATACCG 4677
QY 1262 GCACCTTCGATTCGGTATAGATGCAACACGTGTGTCAACCGAGAGCTCGACTTCAGCC 1321
Db 4678 GCACCTTCGATTCGGTATAGATGCAACACGTGTGTCAACCGAGAGCTCGACTTCAGCC 4737
QY 1322 TTGACCTTACTTTCACCATTTAGACAATCAAGCTTCCCGAGGATGCTGTCTCCGCTACTC 1381
Db 4738 TTGACCTTACTTTCACCATTTAGACAATCAAGCTTCCCGAGGATGCTGTCTCCGCTACTC 4797
QY 1382 AAGTCCGGGTAGGACTGGCAGAGGAGCCAGGATCTACAGATTTGTGGCACCAGGGG 1441
Db 4798 AAGTCCGGGTAGGACTGGCAGAGGAGCCAGGATCTACAGATTTGTGGCACCAGGGG 4857

QY 1442 AGCGTCTTCTGCGCATGTTTGAATCTGCTCTCTGAGTGTATGACGCGGGTGTG 1501
Db 4858 AGCGCCCTCTCGGCATGTTTGAATCTGCTCTCTGAGTGTATGACGCGGGTGTG 4917
QY 1502 CTTGATATGACTTTACGCCCCGAGACCAAGTTAGGCTACGAGCATATATGAACACCC 1561
Db 4918 CTTGATATGACTTTACGCCCCGAGACTACAGTTAGGCTACGAGCATATGAACACCC 4977
QY 1562 CGGACTTCCCGTGTGCCAGAACCATCTTGAATTTTGGAGGGGCTCTTTTACGGCTTCA 1621
Db 4978 CGGGGCTTCCCGTGTGCCAGAACCATCTTGAATTTTGGAGGGGCTCTTTTACGGCTTCA 5037
QY 1622 CCCACATAGACGCCCCACTTCTTATCCCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 1681
Db 5038 CTATATATAGATGCCCATCTTCTATCCCCAGACAAAGCAGAGTGGGAGAACCTTCCCTTACC 5097
QY 1682 TGGTAGCGTACAAAGCCACCGTGTGCGCTAGAGCTCAAGCCCCCTCCCGCTCGTGGGACC 1741
Db 5098 TGGTAGCGTACAAAGCCACCGTGTGCGCTAGGGCTCAAGCCCCCTCCCGCTCGTGGGACC 5157
QY 1742 AGATGTGGAAGTGTGATCGCTCTCAAGCCCCCTCCATGGGCGCAACACTCTGTCTAT 1801
Db 5158 AGATGTGGAAGTGTGATCGCTCTCAAGCCCCCTCCATGGGCGCAACACTCTGTCTAT 5217
QY 1802 ATAGACTGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATATCA 1861
Db 5218 ACAGACTGGGCGCTGTTCAGAAATGAATCAACCTGAGCACCAGTCAACCAATATATCA 5277
QY 1862 TGACATGATGTGGGCTGACCTGGAGTGTCTACAGAGTACCTGGTGTCTCGTGGCGGG 1921
Db 5278 TGACATGATGTGGGCGGACCTGGAGTGTCTACAGAGTGTCTCGTGGCGGG 5337
QY 1922 TTCTGGCTGTTCGGCGGCTATTGCCCTATCCAGAGTGTCTCATAGTAGGTAGGA 1981
Db 5338 TCCTGGCTGTTCGGCGGCTATTGCCCTGTCAAGAGTGTCTCATAGTAGGTAGGA 5397
QY 1982 TTGCTCTTCTCGGGAAGCGGCAATCATACCCGACAGGAAAGTCTCTACCGGGAGTTG 2041
Db 5398 TCGTCTTCTCGGGAAGCGGCAATCATACCTGACAGGAAAGTCTCTACCGAGAGTTG 5457
QY 2042 ATGAATGGAAGTGTCT 2059
Db 5458 ATGAGATGGAAGTGTCT 5475

RESULT.15

AAT99981
ID AAT99981 standard; DNA; 9401 BP.

XX AC AAT99981;
XX 25-MAR-2003 (revised)
DT 16-MAR-1998 (first entry)

XX HCV polyprotein coding sequence.

XX PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
KW C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
KW NS4; DB.

XX Hepatitis C virus.

XX Key Location/Qualifiers
FH 342..9377
FT /*tag= a

XX US5683864-A.

XX 04-NOV-1997.

XX 07-JUL-1992; 92US-00910760.

XX

PR 18-NOV-1987; 87US-00122714.
 PR 30-DEC-1987; 87US-00139886.
 PR 26-FEB-1988; 88US-00161072.
 PR 06-MAY-1988; 88US-00191263.
 PR 26-OCT-1988; 88US-00263584.
 PR 14-NOV-1988; 88US-00271450.
 PR 17-MAR-1989; 89US-00325338.
 PR 20-APR-1989; 89US-00341334.
 PR 21-APR-1989; 89US-00353896.
 PR 18-MAY-1989; 89US-00355002.
 PR 04-APR-1990; 90US-00504352.
 XX (CHIR) CHIRON CORP.
 XX
 XX Kuo G, Houghton M, Choo Q;
 XX
 XX WPI: 1997-548976/50.
 DR P-PSDB; AAW34480.
 XX
 PT Combination of three hepatitis C virus antigens - used for detection of
 PT specific antibodies to diagnose infection.
 XX
 PS Disclosure; Col 25-46; 57pp; English.
 XX
 CC This sequence represents the Hepatitis C virus polypeptide coding
 CC sequence. Fragments of this sequence can be amplified and used in the
 CC combination of HCV antigens of the invention. The HCV antigen combination
 CC comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)
 CC 1-120 of the HCV polypeptide), or its immunologically reactive fragment
 CC containing at least 8 aa. It also comprises two additional antigens from
 CC two different polypeptide domains, including at least 8 aa from the NS3,
 CC NS4, S or NS5 domains of the polypeptide, corresponding, respectively, to
 CC aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polypeptide.
 CC Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa
 CC regions of the HCV polypeptide. These antigen combinations are used
 CC diagnostically to detect anti-HCV antibodies, using any standard
 CC immunoassay format. These antigen combinations have a broader range of
 CC reactivity with antibodies than any antigen individually. (Updated on 25-
 CC MAR-2003 to correct PR field.)
 XX
 SQ Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;
 Query Match 86.7%; Score 1786; DB 2; Length 9401;
 Best Local Similarity 91.7%; Pred. No. 0;
 Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
 2 TGCGCGCTATCAGCGCTATGCCAGCAGACAGAGGGCGCTTTGGGATGCATTAATCACCA 61
 3418 TGCGCGCGCTATCAGCGCGTACGCCAGCAGACAGAGGGCGCTTCTAGGGTGATTAATCACCA 3477
 62 GCTTGTACCGCGCGGACAAAACAGGTGGAGGGTGAGGTTTCAGATCGTCAACTGCTG 121
 3478 GCCTTAAGTGGCGGGACAAAACAGGTGGAGGGTGAGGTTTCAGATCGTCAACTGCTG 3537
 122 CCAGACTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGAGTGTCTTACCATGAGAGCGG 181
 3538 CCCAAACCTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGAGTGTCTTACCATGAGAGCGG 3597
 182 GACAGAGGACCATTCGCTCACTTAAGGTCTCTGTTATCCAGATGTACCAATGTGGAC 241
 3598 GAACAGAGGACCATTCGCTCACTTAAGGTCTCTGTTATCCAGATGTACCAATGTAGAC 3657
 242 AAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGACATTTGG 301
 3658 AAGACCTTGTGGCTGGCGCGCTCCGCAAGGTAGCGCTCATTAACACCATGACATTTGG 3717
 302 GCTCTCGGACCTTTACCTGGTCAAGGACAGCGCGCATGTCATCTGTGCGCGGACGGG 361
 3718 GCTCTCGGACCTTTACCTGGTCAAGGACAGCGCGCATGTCATCTGTGCGCGGACGGG 3777
 362 GTGATGCGAGGCGGAGGCTGCTTTCGCGCGGCTATCTTACTTTGAAGGCTCTCGG 421
 3778 GTGATGCGAGGCGGAGGCTGCTTTCGCGCGGCTATCTTACTTTGAAGGCTCTCGG 3837

QY 422 GAGGCCCTCTGTGTGTCGCCCGCAGGACATGCCGTAGGCATATTCAGAGCCGCGGTATGCA 481
 DB 3838 GGGGTCCGCTGTGTGTGCCCGCGGCGCACGCCGTGGGCATATTTAGGGCCGCGGTGTGCA 3897
 QY 482 CCGGTGGAGTGGCTAAGCGGTGGACTTCATCCCGTAGAGACTTTAGAGACACCATGA 541
 DB 3898 CCGGTGGAGTGGCTAAGCGGTGGACTTTTATCCCTGTGAGAACCTTAGAGACACCATGA 3957
 QY 542 GGTCCCGGTGTCTCAGACAACTCTCCACCACAGCAGTGCCTCCAGAGACTACCAAGTGG 601
 DB 3958 GGTCCCGGTGTCTCAGGATACTCTCTCCAGCAGTAGTGTGCCAGAGCTTCCAGGTGG 4017
 QY 602 CCCACCTGCATGCTCCACCCGAGCGGTAAGAGCACCAGGTGCCGCGGCATACGCAG 661
 DB 4018 CTCACCTCATGCTCCACAGCGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCAG 4077
 QY 662 CTCAGGCTACAGGTGCTGCTCAACCCCTCCGTGCTGCTGCMCAATGGGCTTTGGTG 721
 DB 4078 CTCAGGCTATAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGCTTTGGTG 4137
 QY 722 CTTACATGTCCAAGGCCCATGGATTGATCTTAACATCAGGACTGGGGTGAGACAAATTA 781
 DB 4138 CTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGACAAATTA 4197
 QY 782 CTACTGGCAGCCGATCAGCTATTCCACTACGCAAGTTCCTTGCAGCGCGGTGTT 841
 DB 4198 CCACTGGCAGCCCATCAGTACTCCACCTACGCAAGTTCCTTGCAGCGCGGTGTT 4257
 QY 842 CAGGGGTGCTTATGACATAAATTTGTGAGAGTGCACCTCCAGGATGCAACATCCA 901
 DB 4258 CGGGGGCGCTTATGACATAAATTTGTGAGAGTGCACCTCCAGGATGCAACATCCA 4317
 QY 902 TCTTGGGCTTGGCAGTCTGCTTTCAGCAAGAGAGACCGCGGGGGAGAGTGCAGTGTGC 961
 DB 4318 TCTTGGGCTTGGCAGTCTGCTTTCAGCAAGAGAGTGCAGGGGGAGAGTGCAGTGTGC 4377
 QY 962 TCGCCACCGCTACCCCTCCGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGTTG 1021
 DB 4378 TCGCCACCGCCACCCCTCCGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGTTG 4437
 QY 1022 CTCTGTCACTACCGAGAGATCCCTTTTATGGCAGGCTATTCCTTGAAGCAATTA 1081
 DB 4438 CTCTGTCACTACCGAGAGATCCCTTTTATGGCAGGCTATTCCTTGAAGCAATTA 4497
 QY 1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAGAGTGCAGAGTGCAGCTCGCGCAA 1141
 DB 4498 AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAGAGTGCAGAGTGCAGCTCGCGCAA 4557
 QY 1142 AACTGTGCGGTGTGGGCGTCAATGTGCGGTGCTTACTACCGCGGCTTGTATGTCCGTCA 1201
 DB 4558 AGTGTGCGCATTTGGGCAATCAATGTGCGGTGCTTACTACCGGCTTGTATGTCCGTCA 4617
 QY 1202 TCCGACAGTGGTGTGCTGCTGCGCAACTGACGCCCTCATGACGGGCTTTACCG 1261
 DB 4618 TCCGACAGTGGTGTGCTGCTGCGCAACTGACGCCCTCATGACGGGCTTTACCG 4677
 QY 1262 GCGACTTCGATTCGGGTGATGACTGCAACAGTGTGTCAACCGAGAGTGCAGTTCAGGCC 1321
 DB 4678 GCGACTTCGATTCGGGTGATGACTGCAACAGTGTGTCAACCGAGAGTGCAGTTCAGGCC 4737
 QY 1322 TTGACCTTACCTTACCATTTAGACAAATCACGCTTCCCGAGGATGCTGTCTCCGCTCTC 1381
 DB 4738 TTGACCTTACCTTACCATTTAGACAAATCACGCTTCCCGAGGATGCTGTCTCCGCTCTC 4797
 QY 1382 AACGTGGGGTGGAGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1441
 DB 4798 AACGTGGGGTGGAGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4857
 QY 1442 AGCGTCTTCTGCGCATGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
 DB 4858 AGCGTCTTCTGCGCATGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4917

Mon Feb 28 11:40:44 2005

QY	1502	CTTGGTATGAGTTAGCCCGCCGAGACACAGTTAGGCTAGAGCATACATGAACACC	1561
Db	4918	CTTGGTATGAGTTAGCCCGCCGAGACTACAGTTAGGCTAGAGCGTACATGAACACC	4977
QY	1562	CGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGTCTTTACGGGTCTCA	1621
Db	4978	CGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGTCTTTACAGGCCTCA	5037
QY	1622	CCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGGAAACCTTCCCTATC	1681
Db	5038	CTCATATAGATGCCCACTTCTTATCCAGACAAAGCAGAGTGGGGAGACCTTCCCTTACC	5097
QY	1682	TGGTAGCGTACCAAGCCACCGTGTGCGTGTAGAGCTCAAGCCCTTCCCGTGTGGGACC	1741
Db	5098	TGGTAGCGTACCAAGCCACCGTGTGCGTGTAGGCTCAAGCCCTTCCCGTGTGGGACC	5157
QY	1742	AGATGTGGAAGTCTTGATCCGCTCAAGCCACCCCTCCATGGGCCACACCTCTGCTAT	1801
Db	5158	AGATGTGGAAGTCTTGATTCGCTCAAGCCACCCCTCCATGGGCCACACCCCTGCTAT	5217
QY	1802	ATAGACTGGGCGCTGTCCAGAATGAAGTCAACCTGACGCCACCCAGTCAACAAATATCA	1861
Db	5218	ACAGACTGGGCGCTGTTCAGATGAATCAACCTGACGCCACCCAGTCAACAAATATCA	5277
QY	1862	TGACATGTATGTGCGGTGACCTGGAGGTGCTACAGAGTACCTGGGTGCTCGTTGGGGCG	1921
Db	5278	TGACATGTATGTGCGGCGACCTGGAGGTGCTACAGAGTACCTGGGTGCTCGTTGGGGCG	5337
QY	1922	TTCTGGCTGCTTGGGCGGTATTGCTATCCACAGGCTGCGTGTGCTATAGTAGGTAGGA	1981
Db	5338	TCCTGGCTGCTTGGGCGGTATTGCTGTCAACAGGCTGCGTGTGCTATAGTAGGTAGGA	5397
QY	1982	TTGTCTTGTCCGGAAGCCGGCAATCATACCGACAGGGAAGTCTCTTACCGGGAGTTCCG	2041
Db	5398	TCGTCTTGTCCGGAAGCCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGATTCCG	5457
QY	2042	ATGAATGGAAGATGCT	2059
Db	5458	ATGAGATGGAAGATGCT	5475

Search completed: February 27, 2005, 18:30:06
Job time : 1064 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 17:06:13 ; Search time 370 Seconds
(without alignments)
9114.504 Million cell updates/sec

Title: US-09-930-591-1
Perfect score: 2061
Sequence: 1 atggcgctatcacggccta.....atgaatggaagagtgtga 2061

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1786	86.7	7310	3	US-08-444-818-74
2	1786	86.7	9379	3	US-09-388-874-1
3	1786	86.7	9379	4	US-09-916-359-1
4	1786	86.7	9401	1	US-07-910-760-9
5	1786	86.7	9401	1	US-08-440-519-9
6	1786	86.7	9401	3	US-08-440-549-9
7	1786	86.7	9401	3	US-08-823-895A-25
8	1785.2	86.6	9401	5	PCT-US91-02225-9
9	1784.4	86.6	6785	3	US-08-444-818-65
10	1784.4	86.6	8316	3	US-08-444-818-88
11	1784.4	86.6	8987	3	US-08-444-818-137
12	1784.4	86.6	9185	3	US-08-444-818-122
13	1784.4	86.6	9185	3	US-08-444-818-123
14	1784.4	86.6	9379	3	US-08-444-818-176
15	1781.2	86.4	2058	4	US-09-881-239-2
16	1781.2	86.4	5360	3	US-08-444-818-53
17	1779.6	86.3	2058	4	US-09-881-654-1
18	1779.6	86.3	2058	4	US-10-637-323-1
19	1774.8	86.1	12980	3	US-08-811-566-5
20	1774.8	86.1	12980	3	US-09-034-756-5
21	1773.2	86.0	9646	3	US-08-811-566-1
22	1773.2	86.0	9646	3	US-09-034-756-1
23	1765.2	85.6	9599	3	US-09-014-416-2
24	1765.2	85.6	9599	3	US-09-014-416-6
25	1762	85.5	9416	3	US-08-823-895A-26
26	1762	85.5	9416	4	US-10-104-966-13
27	1750.8	84.9	9401	2	US-08-432-693-1

28	1750.8	84.9	9416	3	US-08-811-566-19	Sequence 19, Appl
29	1750.8	84.9	9416	3	US-09-034-756-19	Sequence 19, Appl
30	1749.2	84.9	9365	4	US-09-827-688-7	Sequence 7, Appl
31	1557.8	75.6	2064	1	US-08-350-884-69	Sequence 69, Appl
32	1557.8	75.6	2064	1	US-08-440-548-69	Sequence 69, Appl
33	1557.8	75.6	2064	1	US-08-709-173-69	Sequence 69, Appl
34	1557.8	75.6	2064	2	US-08-709-177-69	Sequence 69, Appl
35	1557.8	75.6	2523	1	US-08-350-884-85	Sequence 85, Appl
36	1557.8	75.6	2523	1	US-08-440-548-85	Sequence 85, Appl
37	1557.8	75.6	2523	1	US-08-709-173-85	Sequence 85, Appl
38	1557.8	75.6	2523	2	US-08-709-177-85	Sequence 85, Appl
39	1546	75.0	1932	3	US-09-128-314-1	Sequence 1, Appl
40	1546	75.0	8157	3	US-09-128-314-3	Sequence 3, Appl
41	1398.4	67.9	9413	4	US-09-827-688-6	Sequence 6, Appl
42	1381.2	67.0	7989	4	US-09-539-601-10	Sequence 10, Appl
43	1381.2	67.0	8001	4	US-09-539-601-7	Sequence 7, Appl
44	1379.6	66.9	8001	4	US-09-539-601-22	Sequence 22, Appl
45	1379.2	66.9	8638	4	US-10-029-907-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-444-818-74
; Sequence 74, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..7310
; US-08-444-818-74

Query Match 86.7%; Score 1786; DB 3; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGCGCGCTATCACGGCTATGCCAGCAGACAAGGGCCCTTTTGGATGTCATAATCACCA 61
DB 1729 TGGCGCCCATACAGGGTACGCCAGCAGACAAGGGCCCTCTAGGTGTCATATACCA 1788
QY 62 GCTTGAACGGCGGGGACAAAACCAAGGTGGAGGGTGAGGTTTCAGATCGTGTCAACTGCTG 121
DB 1789 GCTTAACATGCGCGGACAAAACCAAGTGGAGGGTGAGGTTCAGATTTGTCAACTGCTG 1848
QY 122 CCCAGACTTTCTTGGCAACCTGCATTAACGGGGTGTGTGAGTGTCTTACCAATGAGCGG 181
DB 1849 CCCAAACCTTCTGCGCAACGTGCATCAATGGGGTGTGTGAGTGTCTTACCAACGGGGCG 1908
QY 182 GAAACAAGGACCAATGCGTCACCTAAGGGTCTGTATCCAGATGTACACCAATGTGGACC 241
DB 1909 GAAACGAGGACCAATGCGGTACCAAGGGTCTGTATCCAGATGTATACCAATGTAGACC 1968
QY 242 AAGACCTCTAGGCTGGCCCGCTGCCCAAGGTGCCCGTCAATTAACACCAATGCACTTGG 301
DB 1969 AAGACCTTGTGGCTGGCCCGCTGCCCAAGGTAGCGCTCAATTAACACCTGCACTTGG 2028
QY 302 GCTCTCGACCTTTACCTGTGTACAGGSCACGCCGATGTCAATCTGTGTGCGCCGACGGG 361
DB 2029 GCTCTCGACCTTTACCTGTGTACAGGSCACGCCGATGTCAATCTGTGTGCGCCGACGG 2088
QY 362 GTATGSGCAGGGGACGCTCTTTCGCCCGGCTATCTTACTTTGAAAGGCTCCTCGG 421
DB 2089 GTATGAGAGGGGACGCTCTGTGCGCCCGGCTATCTTACTTTGAAAGGCTCCTCGG 2148
QY 422 GAGGCGCTCTGTGTGCCCGGACGACATGCCGTAGGCATATTCAGAGCGCGGTATGCA 481
DB 2149 GGGTCCGCTGTGTGCCCGGCGGACGCGCTGGGCATATTTAGGGCGCGGTGCA 2208
QY 482 CCCGTGAGTGTGCTAAGGCGGTGAGTTCATCTCCCGTAGAGAGCTTAGAGACAACATGA 541
DB 2209 CCCGTGAGTGTGCTAAGGCGGTGAGTTCATCTCTCCCGTAGAGAGCTTAGAGACAAC 2268
QY 542 GGTCCCGGTGTCTCAGACAACCTCTCCCGACAGCAGTCCCGACAGCTACCAAGTGG 601
DB 2269 GGTCCCGGTGTCTCAGGATACTCTCTCCAGTAGTGTCCCGACAGCTTCAGAGTGG 2328
QY 602 CCCACCTGATGCTCCACCGCAGCGTAAGAGCACCAAGTTCGCGCGCGATACGACG 661
DB 2329 CTCACCTCATGCTCCACAGCAGCGCAAAAGCACCAAGTTCGCGGTGCATATGCA 2388
QY 662 CTCAGGGCTACAGGTGCTGTGCTCAACCCCTCGTGTCTGCAACAATGGGCTTTGGT 721
DB 2389 CTCAGGGCTAAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGCTTTGGT 2448
QY 722 CTTTACATGTCCAAGGCCATGGATGATCTTAACATCAGGACTGGGGTGAGGACAATTA 781
DB 2449 CTTTACATGTCCAAGGCTCATGGATGATCTTAACATCAGGACTGGGGTGAGGACAATTA 2508
QY 782 CTAATGGCAGCCGATACGATATTCACCTACGGCAAGTTCCTTCCGACGGCGGTGT 841
DB 2509 CCACATGGCAGCCCATCAGTACTCCACCTACGGCAAGTTCCTTCCGACGGCGGTGT 2568
QY 842 CAGGGGCTCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAATCA 901
DB 2569 CAGGGGCGCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAATCA 2628
QY 902 TCTTTGGGCTATGGCATCTGTCTTACCAAGCAGAGACCGCGGGGCGAGACTGACTGTGC 961
DB 2629 TCTTTGGGCTATGGCATCTGTCTTACCAAGCAGAGACTCGCGGGGCGAGACTGACTGTGC 2688
QY 962 TCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGTTG 1021
DB 2689 TCGCCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGTTG 2748
QY 1022 CTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATCCCTTTGAGCAATTA 1081
DB 2749 CTCTGTCCACCGAGAGATCCCTTTTATGGCAAGGCTATCCCTTTGAGCAATTA 2808

QY 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAAGTGGACAGCTCGCGCAA 1141
DB 2809 AGGGGGGAGACATCTCATCTTCTGTCTTCAAGAGAAGTGGACAGCTCGCGCAA 2868
QY 1142 AACTGGTCGGTGTGGCGCTCAATGCGGTGCTTACTACGGCGCTTCAATGTGTCCGTCA 1201
DB 2869 AGCTGGTCGCATTTGGGCATCAATGCGGTGCTTACTACGGCGCTTCAATGTGTCCGTCA 2928
QY 1202 TCCGACCAAGTGTGACGTTGTCGTCGTCGCAACTGACGCGCTCATGACCGCTTTACCG 1261
DB 2929 TCCGACCAAGCGGCGATGTTGTGTCGTCGTCGCAACCGATGCCCTCATGACCGCTTACCG 2988
QY 1262 GCGACTTCGATTCGGTGATAGACTGCAACACGTTGTCCACAGACAGTTCGATTCAGCC 1321
DB 2989 GCGACTTCGATTCGGTGATAGACTGCAATACGTTGTCCACAGACAGTTCGATTTACGCC 3048
QY 1322 TTGACCTTACCTTACCAATTGAGACAATCAAGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381
DB 3049 TTGACCTTACCTTACCAATTGAGACAATCAAGCTTCCCGTCCCGCAGGATGCTGTCTCCCGACTC 3108
QY 1382 AACGTCGGGGTAGGACTGGCAGAGGGAAGCAGGATCTACAGATTTTGTGGCACCGGGG 1441
DB 3109 AACGTCGGGGCAGGACTGGCAGGGGAAGCAGGATCTACAGATTTTGTGGCACCGGGG 3168
QY 1442 AGCGTCTTCTGGCATGTTTGAATCTGCTCTCTGCGAGTGTATGACGCGGTTGTG 1501
DB 3169 AGCGTCTTCTGGCATGTTTGAATCTGCTCTCTGCGTCTCTGAGTGTATGACGCGGTTGTG 3228
QY 1502 CTTGTATGAGTTACGCGCGCGAGACCAAGTTAGCTAGGAGTACGATGATGAAACGCC 1561
DB 3229 CTTGTATGAGTTACGCGCGCGAGACCAAGTTAGCTAGGAGTACGATGATGAAACGCC 3288
QY 1562 CGGACTTTCCTGTCGCAAGACCATCTTGAATTTTGGAGGGCTCTTTACGGGTCTCA 1621
DB 3289 CGGACTTTCCTGTCGCAAGACCATCTTGAATTTTGGAGGGCTCTTTACAGGCTCTCA 3348
QY 1622 CCCACATAGACGCCCATCTTCTATCCACAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
DB 3349 CTCATATAGTCCCATCTTCTATCCACAGACAAAGCAGAGTGGGAGAACCTTCTTACC 3408
QY 1682 TGGTAGCGTACAGGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
DB 3409 TGGTAGCGTACAGGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 3468
QY 1742 AGATGTGAAGTGTGTTGATTCGCTCAAGCCACCCCTCCATGGGCAACACCTCTGTCTAT 1801
DB 3469 AGATGTGAAGTGTGTTGATTCGCTCAAGCCACCCCTCCATGGGCAACACCCCTGTCTAT 3528
QY 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAGTATATCA 1861
DB 3529 ACAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAGTATATCA 3588
QY 1862 TGACATGTATGTGCGCTGACCTGGAGGTGCTCAAGAGTACCTGGGTGCTCGTTGGGGGG 1921
DB 3589 TGACATGTATGTGCGGCGACCTGGAGGTGCTCAAGAGTACCTGGGTGCTCGTTGGGGGG 3648
QY 1922 TTTCTGGTGTCTTGGCGCGTATTCCTTATCCAGAGGTGCGGTGCTCATAGTAGTAGGA 1981
DB 3649 TCTTGGTGTCTTGGCGCGTATTCCTTATCCAGAGGTGCGGTGCTCATAGTAGTAGGA 3708
QY 1982 TTGCTTGTTCGGAAGCGCGCAATCATACCGACAGGGAAGTCTCTTACCGGGAGTTG 2041
DB 3709 TCGTCTTGTTCGGAAGCGCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTG 3768
QY 2042 ATGAATGGAAGTGTCT 2059
DB 3769 ATGAGATGGAAGTGTCT 3786

RESULT 2
US-09-388-874-1
; Sequence 1, Application US/09388874
; Patent No. 6284249


```
QY 1802 ATAGACTGGCGCTGTCAGAAATAGTCAACCTGACGACCCAGTCAACAAAGTATATCA 1861
Db 5196 ACAGACTGGCGCTGTCAGAAATAGTCAACCTGACGACCCAGTCAACAAATATCA 5255
QY 1862 TGACATGATGTCGGCTGACCTGGAGGTGCTCAAGATACCTGGGTGCTGTTGGCGGG 1921
Db 5256 TGACATGATGTCGGCGGACCTGGAGGTGCTCAAGATACCTGGGTGCTGTTGGCGGG 5315
QY 1922 TTCTGCTGCTTTGGCGCGGTATTCCTATCCACAGGCTGCTGCTCATAGTAGTAGGA 1981
Db 5316 TCTTGCTGCTTTGGCGCGGTATTCCTATCCACAGGCTGCTGCTCATAGTAGTAGGA 5375
QY 1982 TTGCTTTGTCGGAAGCGGCAATCATACCGGACAGGAAAGTCTCTACCGGAGTTGG 2041
Db 5376 TCGTCTGTCGGAAGCGGCAATCATACCTGACAGGAAAGTCTCTACCGAGATTGG 5435
QY 2042 ATGAATGAAGAGTGCT 2059
Db 5436 ATGAGATGAAGAGTGCT 5453

RESULT 3
US-09-916-359-1
; Sequence 1, Application US/09916359
; Patent No. 6538123
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; TITLE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: PMCF97-03A
; CURRENT APPLICATION NUMBER: US/09/916,359
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 09/388,874
; PRIOR FILING DATE: 1993-09-02
; PRIOR APPLICATION NUMBER: 97/02,887
; PRIOR FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320)...(9352)
US-09-916-359-1

Query Match 86.7%; Score 1786; DB 4; Length 9379;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTATACGGCTATGCCAGCAGACAAAGGGGCTTTGGGATGCATAATCACCA 61
Db 3396 TGGCGCTATACGGCTATGCCAGCAGACAAAGGGGCTTTGGGATGCATAATCACCA 3455
QY 62 GCTTGACCGGCGGACAAAACAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 121
Db 3456 GCTTGACCGGCGGACAAAACAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 3515
QY 122 CCCAGACTTTTGGCAACTGCATTAAACGGGGTGTGTTGGACTGTCTACCAATGGAGCG 181
Db 3516 CCCAAACCTTCTGGCAACTGCATTAAACGGGGTGTGTTGGACTGTCTACCAATGGAGCG 3575
QY 182 GAACAGGACCAATGGGTGCTTACCTAGGCTGCTGTTATCCAGATGTACCAATGGAGCC 241
Db 3576 GAACAGGACCAATGGGTGCTTACCTAGGCTGCTGTTATCCAGATGTACCAATGGAGCC 3635
QY 242 AAGACTCTGAGCTGGCGCGCTCCCAAGGTGCCGCTCATTAAACCACTGCTTGG 301
Db 3636 AAGACTCTGAGCTGGCGCGCTCCCAAGGTGCCGCTCATTAAACCACTGCTTGG 3695
QY 302 GCTCTCGGACCTTTACCTGGTCAAGGACACCGGATGTCTATCTGTGCGCGGCGG 361
```

```
Db 3696 GCTCTCTCGGACCTTTACCTGGTCAAGAGGACCGCGATGTCTATCCCTGTCGCGCGGG 3755
QY 362 GTGATGGCAGGGGACGCTGCTTTCCGCCCGGCTATCTCTTACTTTGAAGGCTCTCGG 421
Db 3756 GTGATGGCAGGGGACGCTGCTTTCCGCCCGGCTATCTCTTACTTTGAAGGCTCTCGG 3815
QY 422 GAGGCGCTCTGTGTCGCCCGGACGACATGCGGTAGGCAATTTACAGAGCGCGGTATGA 481
Db 3816 GGGGTCCGCTGTTGTCGCCCGGCGGACGCGGTGGGCAATTTAGGGCGCGGTATGA 3875
QY 482 CCGGTGGAGTGGCTAAGCGGTGAGTCTATCCCTCGTAGAGCTTTAGAGCAACATGA 541
Db 3876 CCGGTGGAGTGGCTAAGCGGTGAGTCTATCCCTCGTAGAGCAACCTAGAGCAACATGA 3935
QY 542 GGTCCCCCGGTGTTCTCAGACAACTCTCCACAGCAGTGCCTCCAGAGCTTACCAAGTG 601
Db 3936 GGTCCCCCGGTGTTCTCAGAGTAACTCTCTCCACAGTGTGCTCCAGAGCTTCCAGGTGG 3995
QY 602 CCCACTGCTGCTCCCAAGCGGCTAAGAGCAACCAAGGTCCCGGCGCATACGCGAG 661
Db 3996 CTCACCTCCATGCTCCCAAGCGGCTAAGAGCAACCAAGGTCCCGGCGCATATGCGAG 4055
QY 662 CTCAGGCTCAAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 721
Db 4056 CTCAGGCTCAAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 4115
QY 722 CTTACATGCTCAAGGCGGCTGAGTGTGATCTTAACATCAGGACTGGGTGAGGCAATTA 781
Db 4116 CTTACATGCTCAAGGCGGCTGAGTGTGATCTTAACATCAGGACTGGGTGAGGCAATTA 4175
QY 782 TCTTGGCGAGCCGCTACAGTATTCACCTACGCGCAAGTTCCTTCCCGACGCGGCTGTT 841
Db 4176 CCACTGGCAGCCGCTACAGTATTCACCTACGCGCAAGTTCCTTCCCGACGCGGCTGTT 4235
QY 842 CAGGGGTGCTTATGACATAATAATTTGTGAGTGTGCTTCAACGAGTGTGCAACATCA 901
Db 4236 CAGGGGTGCTTATGACATAATAATTTGTGAGTGTGCTTCAACGAGTGTGCAACATCA 4295
QY 902 TCTTGGCGATTTGGCACTGCTTCCAGCAGAGACCGCGGGGCGAGACTGCTGTCG 961
Db 4296 TCTTGGCGATTTGGCACTGCTTCCAGCAGAGACTGCGGGGCGAGACTGCTGTCG 4355
QY 962 TCGCCACCGCTACCCCTCGGCTCCGTCACCTGTCGCCCATCTTAACATTCAGAGGTTG 1021
Db 4356 TCGCCACCGCTACCCCTCGGCTCCGTCACCTGTCGCCCATCTTAACATTCAGAGGTTG 4415
QY 1022 CTTGTCCTACCTACCGGAGAGTCCCTTTTATGGAAGGCTATTCCTTGAAGCAATTA 1081
Db 4416 CTTGTCCTACCTACCGGAGAGTCCCTTTTATGGAAGGCTATTCCTTGAAGCAATTA 4475
QY 1082 AGGGGGGAGACATCTCATCTTCTCCACTCAAGAAGAGTGCAGAGCTCCGCGCA 1141
Db 4476 AGGGGGGAGACATCTCATCTTCTCCACTCAAGAAGAGTGCAGAGCTCCGCGCA 4535
QY 1142 AACTGCTCGCTGCGGCTCAATGCGGTGCTTACTACCGGCGCTTCAATGTGCTGCTCA 1201
Db 4536 AGCTGCTCGCTGCGGCTCAATGCGGTGCTTACTACCGGCGCTTCAATGTGCTGCTCA 4595
QY 1202 TCCCGACAGTGGTGAAGTGTGTCGTCGTCGCAACTGACGCGCTCATGACCGGCTTACCG 1261
Db 4596 TCCCGACAGTGGTGAAGTGTGTCGTCGTCGCAACTGACGCGCTCATGACCGGCTTACCG 4655
QY 1262 GCGACTTCGATTCGCTGATAGCTGCAACAGTGTGTCACAGAGTGCAGAGCTTCAGCC 1321
Db 4656 GCGACTTCGATTCGCTGATAGCTGCAACAGTGTGTCACAGAGTGCAGAGCTTCAGCC 4715
QY 1322 TTGACCTTACCTTACCAATTCAGCAATTCAGGCTTTCCCGAGGATGCTGTCTCCGCTACTC 1381
Db 4716 TTGACCTTACCTTACCAATTCAGCAATTCAGGCTTTCCCGAGGATGCTGTCTCCGCTACTC 4775
QY 1382 AACGTCGGGTGAGTGTGCTGCGAGAGGAAAGCAGGATCTACAGATTTGTGGACCGGGGG 1441
Db 4776 AACGTCGGGTGAGTGTGCTGCGAGAGGAAAGCAGGATCTACAGATTTGTGGACCGGGGG 4835
```


[illegible]


```

;
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4167
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Leu."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4323
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4701
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Tyr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4752
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ser."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5970
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Gly."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6183
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be His."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6186
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Cys."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6402
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7386
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ser."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7494
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Phe."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7497
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ala."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7845
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Phe."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8409
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Gly."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9102
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also by Gly."
; FEATURE:
; NAME/KEY: misc_feature

```

```

;
; LOCATION: 9327
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Pro."
; US-08-440-519-9
;
; Query Match
; Best Local Similarity 86.7%; Score 1786; DB 1; Length 9401;
; Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
;
; QY 2 TGGCGCCTATCAGCGCCTATGCCAGCAGACAGAGGGGCTTTTGGGATGCATAATCACCA 61
; Db 3418 TGGCGCCCATCAGCGCGTACGCCAGCAGACAGAGGGGCTCTCTAGGGTGCATAATCACCA 3477
;
; QY 62 GCTTGACCGCGCGGACAAACACAGGTGGAGGTGAGGTTTCAGATCTGTCACTGTG 121
; Db 3478 GCCTAACTGGCGCGGACAAACCAAGTGGAGGTGAGGTCAGATTGTGTCACCTGTG 3537
;
; QY 122 CCCAGACTTTCTTGGCAACCTGCATTAAACGGGGTGTGTGGACTGTCTACCATGGAGCG 181
; Db 3538 CCCAAACCTTCTTGGCAACGTGCATCAATGGGTGTCTGGACTGTCTACCAAGGGCG 3597
;
; QY 182 GAAAGGACCATTCGCTACCTTAAGGTCTCTGTATCCAGATGTACACCATGTGGACC 241
; Db 3598 GAAAGGACCATTCGCTACCTTAAGGTCTCTGTATCCAGATGTATACCAATGTAGACC 3657
;
; QY 242 AAGACCTCGTAGGCTGGCCCGCTCCCAAGTCCCGCTCATTAACACCATGCACCTTGG 301
; Db 3658 AAGACCTTGTGGGTGGCCCGCTCCGCAAGGTAGCCGCTCATTTGACACCTGTGCTGG 3717
;
; QY 302 GCTCTCGGACCTTTTACCTGTGTACAGGACGCGCGATGTCTATCTGTGGCCGACGG 361
; Db 3718 GCTCTCGGACCTTTTACCTGTGTACAGGACGCGCGATGTCTATCTGTGGCCGCG 3777
;
; QY 362 GTGATGGCAGGGGCHAGCCTGTCTTTCGCCCGCGGCTATCTCTTACTTTGAAAGCTCTCG 421
; Db 3778 GTGATAGCAGGGGCHAGCCTGTCTTTCGCCCGCGGCTATCTCTTACTTTGAAAGCTCTCG 3837
;
; QY 422 GAGGCCCTCTGTGTGCCCGCAGGACATGCGTAGGCATATTACAGAGCCGGGTATGCA 481
; Db 3838 GGGGTCCGCTGTGTGCGCCCGCGGGGACGCGGTGGGATATTAGGGCCGGGGTGTGCA 3897
;
; QY 482 CCGGTGGAGTGGCTAAGCGGTGGACTTTTCATCCCGTAGAGAGCTTAGAGCAACATGA 541
; Db 3898 CCGGTGGAGTGGCTAAGCGGTGGACTTTTCATCCCGTAGAGAACCTAGAGCAACATGA 3957
;
; QY 542 GGTCCCGGTGTCTCAGACAACTCTCCACACAGAGTCCCGCAGAGCTACCAAGTGG 601
; Db 3958 GGTCCCGGTGTCTCAGGATAACTCTCTCCACAGTAGTCCCGCAGAGCTTCCAGGTGG 4017
;
; QY 602 CCCACCTGCATGCTCCACCGGACGCGTAAGAGCAACCAAGTCCCGCGCATACGCAG 661
; Db 4018 CTCACCTCATGCTCCACAGGACGCGCAAAAGCAACCAAGGTCCCGGTGCATATGCAG 4077
;
; QY 662 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCGTTGCTGCAACAATGGGCTTTGGTG 721
; Db 4078 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCGTTGCTGCAACAATGGGCTTTGGTG 4137
;
; QY 722 CTTACATGTCCAAAGGCCCATGGGATTGATCTTAACATCAGGACTGGGGTGGAGCAATTA 781
; Db 4138 CTTACATGTCCAAAGGCTCATGGATTCGATCTTAACATCAGGACCGGGGTGGAGCAATTA 4197
;
; QY 782 CTACTGGCAGCCCGGATCAGTATTCCACTACGGCAAGTTCCTTCCCGACGGCGGTGT 841
; Db 4198 CCACTGGCAGCCCGGATCAGTATTCCACTACGGCAAGTTCCTTCCCGACGGCGGTGT 4257
;
; QY 842 CAGGGGGTCTTATGACATAATAATTTGTGACAGTGCACCTCCACGAGTGAACATCA 901
; Db 4258 CAGGGGGTCTTATGACATAATAATTTGTGACAGTGCACCTCCACGAGTGAACATCA 4317
;
; QY 902 TCTTGGGCATTTGGCAGTCTCTTGTGACCAAGCAGAGACCGCGGGGCGAGACTGTGTC 961
; Db 4318 TCTTGGGCATTTGGCAGTCTCTTGTGACCAAGCAGAGACTGTGCGGGGCGAGACTGTGTC 4377

```

962 TCDCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCCCATCTTAACATCGAGGAGTTG 1021
Db TCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCCCATCTTAACATCGAGGAGTTG 4437
1022 CTCTGTCCACTACCGGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATTA 1081
Db CTCTGTCCACCACCGGAGATCCCTTTTACGGCAAGGCTATCCCTTCGAAGTAAATCA 4497
1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAAGAGTGCAGAGCTCGCGCAA 1141
Db AGGGGGGAGACATCTCATCTTCTGCACTCAAGAAGAGTGCAGAGCTCGCGCAA 4557
1142 AACTGTTCGGTTGGCGGTCAATGCGGTGCTTACTACCGCGCTTGTATGTCTCGTCA 1201
Db AGTGTGCGCATTTGGGCAATCAATGCGGTGCTTACTACCGCGCTTGTATGTCTCGTCA 4617
1202 TCCGACCACTGTGTGAGTTGTGCTGTGCAACTGACGCCCTCATGACCGGCTTTACCG 1261
Db TCCGACCACTGTGTGAGTTGTGCTGTGCAACTGACGCCCTCATGACCGGCTTTACCG 4677
1262 GCAGCTTCGATTCGGTATAGACTGCAACACGTTGTCTACCCAGACAGTTCAGCC 1321
Db GCAGCTTCGATTCGGTATAGACTGCAACACGTTGTCTACCCAGACAGTTCAGCC 4737
1322 TTGACCTTACCTTCAACATGAGACAAATCAGCTTCCCGAGATGCTGTCTCCGTAATC 1381
Db TTGACCTTACCTTCAACATGAGACAAATCAGCTTCCCGAGATGCTGTCTCCGTAATC 4797
1382 AAGCTCGGGGTAGGACTGCGAGGGAAGCCAGGCACTACAGATTGTGGCACCGGG 1441
Db AAGCTCGGGGTAGGACTGCGAGGGAAGCCAGGCACTACAGATTGTGGCACCGGG 4857
1442 AGCGTCTCTTCTGCACTGTTGACTCTGCTCTGCTGCGAGTGTATGACGGGGTGTG 1501
Db AGCGTCTCTTCTGCACTGTTGACTCTGCTCTGCTGCGAGTGTATGACGGGGTGTG 4917
1502 CTTGGTATAGCTTACCGCGCGGAGACACAGTGTAGGCTACGAGCATCATGAACACC 1561
Db CTTGGTATAGCTTACCGCGCGGAGACTACAGTGTAGGCTACGAGCATCATGAACACC 4977
1562 CGGAGCTTCCGCTGTCAGACATCTTGAATTTGGAGGGGCTTTTACGGTCTCA 1621
Db CGGAGCTTCCGCTGTCAGACATCTTGAATTTGGAGGGGCTTTTACGGGCTCA 5037
1622 CCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 1681
Db CTATATAGATGCCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 5097
1682 TGGTAGCGTACCAAGCCCGTGTGCGTAGAGCTCAAGCCCTTCCCGCTCGTGGACC 1741
Db TGGTAGCGTACCAAGCCCGTGTGCGTAGAGCTCAAGCCCTTCCCGCTCGTGGACC 5157
1742 AGATGTGGAAGTCTTGTATCCGCTCAAGCCACCTTCAATGGGCCAACCTCTGCTAT 1801
Db AGATGTGGAAGTCTTGTATCCGCTCAAGCCACCTTCAATGGGCCAACCTCTGCTAT 5217
1802 ATAGACTGGCGCTGTCCAGAAAGTCACTGACCAAGCAGTCCAGCAAGTATATCA 1861
Db ACAGACTGGCGCTGTTCAGAAATCACTGACCAAGCAGTCCAGCAAGTATATCA 5277
1862 TGACATGTATGTCCGCTGACCTGGAGGTGTCAAGAGTATCTGGGTGCTCGTTGGCGGG 1921
Db TGACATGTATGTCCGCGGACCTGGAGGTGTCAAGAGTATCTGGGTGCTCGTTGGCGGG 5337
1922 TTCTGCGCTGTTGGCGGCTATTTGCTTATCCAGGCTGCGGTGCTATAGTAGTAGGA 1981
Db TTCTGCGCTGTTGGCGGCTATTTGCTTATCCAGGCTGCGGTGCTATAGTAGTAGGA 5397
1982 TTGCTTTGCGGAAGCGGCAATCATACCCAGAGGGAAGTCTTACCGGAGTTCCG 2041
Db TTGCTTTGCGGAAGCGGCAATCATACCTGACAGGGAAGTCTTACCGGAGTTCCG 5457
2042 ATGAAATGGAAGAGTGTCT 2059

Db 5458 ATGAGATGGAAGAGTGTCT 5475

RESULT 6

US-08-440-549-9
; Sequence 9, Application US/08440549
; Patent No. 6312889
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,549
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Bsq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9401 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 342..9374
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 366
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Arg."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 372
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 867
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1341
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ile."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."

FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-08-440-549-9
Query Match 86.7%; Score 1786; DB 3; Length 9401;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 2 TGGCGCCTATCAGCGCCTATGCCAGCAGACAAGGGGCTTTGGGATGCATATCACA 61
DB 3418 TGGCGCCCATCAGCGCGTACGCCAGCAGACAAGGGGCTTCTAGGGTGCATATCACA 3477
QY 62 GCTTGACCGCGGAGCAAAAACAGGTGGAGGGTGAGGTTTCAGATCGTGCAACTGCTG 121
DB 3478 GCCTAACTGGCGGGAGCAAAAACAGGTGGAGGGTGAGGTTTCAGATCGTGCAACTGCTG 3537
QY 122 CCCAGACTTCTTGGCAACCTGCATTAACGGGGGTGTGTGGACTGTCTACCATGGAGCGG 181
DB 3538 CCCAAACCTTCTGTGCAACGTGCATCAATGGGGTGTGTGGACTGTCTTACCACGGGGCGG 3597
QY 182 GAACAAGGACCATTTGCGTCACCTAAGGCTCTGTATCCAGATGTACACCAATGTGGACC 241
DB 3598 GAACGAGGACCATTCGCGTCACCCAAAGGCTCTGTATCCAGATGTATACCAATGTAGACC 3657
QY 242 AAGACTCTGTAGGCTGGCCGCTCCCAAGGTGCCCGCTCATTAACACCATGTCACTTGG 301
DB 3658 AAGACTTGTGGCTGGCCGCTCGCAAGGTAGCGGCTCATTTAGACACCTGTCACTTGG 3717
QY 302 GCTCTCGGACCTTTACCTGTGCAGAGGCGCGGATGTCTTCTGTGGCGGCGAGCGG 361
DB 3718 GCTCTCGGACCTTTACCTGTGCAGAGGCGCGGATGTCTTCTGTGGCGGCGGCGG 3777
QY 362 GTGATGGCAGGGGAGCCTGCTTTTGGCCCGGCGCTATCTTCTTAAAGGCTCCTCGG 421
DB 3778 GTGATAGCAGGGGAGCCTGCTGTGCGCCCGGCGCATTTCTTACTTTAAAGGCTCCTCGG 3837
QY 422 GAGGCCCTCTGCTGTGCCCGCAGGACATGCGGTAGGCATATTTCAGAGCCGCGGTATGCA 481
DB 3838 GGGGTCCGCTGTTGTGCCCGGCGGACGCGGTGGGATATTATTAGGGCGCGGTGTGCA 3897
QY 482 CCGGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCGGTAGAGAGCTTACAGACAACATGA 541
DB 3898 CCGGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCGGTAGAGAGCTTACAGACAACATGA 3957
QY 542 GGTCCCGGCTTCTCAGACAACCTCCCGCAGCAGTCCCGCAGAGCTACCAAGTGG 601
DB 3958 GGTCCCGGCTTCTCAGGATAACTCTCTCCACAGTGTGCGCCAGAGCTTCCAGGTGG 4017
QY 602 CCCACTGCATGCTCCCAACCGGCGGTAGAGCAACCAAGGTCCCGGCGCGCATACGAG 661
DB 4018 CTCACCTCCATGCTCCCAACCGGCGGTAGAGCAACCAAGGTCCCGGCGCATATGAG 4077
QY 662 CTCAGGGCTACAAGGTGTGTGTGTCAACCCCTCGGTGTGTGTGTGTGTGTGTGTGTGT 721
DB 4078 CTCAGGGCTACAAGGT 4137

QY 722 CTTACATGTCCTAAGCCCATGGGATGATCCTTAACATCAGGACTGGGGTGAGSACAATTA 781
DB 4138 CTTACATGTCCTAAGGCTCATGGATCGATCCTTAACATCAGGACCGGGTGAGAACATTA 4197
QY 782 CTTACTGGCAGCCGATCAGTATTCACCTACGGCAAGTTCCTTGGCGACGGGGTGT 841
DB 4198 CCACTGGCAGCCCATCAGTATTCACCTACGGCAAGTTCCTTGGCGACGGGGTGT 4257
QY 842 CAGGGGGTGTATGACATAAATTTGTGACAGTGCACATCCACGGATGCAACATCCA 901
DB 4258 CGGGGGGGCTTATGACATATTAATTTGTGACAGTGCACATCCACGGATGCAACATCCA 4317
QY 902 TCTTGGGCATTTGACATGCTCTTGAACAAAGCAGAGACCGGGGGGCGAGACTGACTGTGC 961
DB 4318 TCTTGGGCATTTGACATGCTCTTGAACAAAGCAGAGACTGCGGGGGGCGAGACTGTTGTGC 4377
QY 962 TCGCACCGCTACCCCTCCGGGCTCGTCACTGTGCGCCCATCTTAACATCGAGGAGTTG 1021
DB 4378 TCGCACCGCCACCCCTCCGGGCTCGTCACTGTGCGCCCATCTTAACATCGAGGAGTTG 4437
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGCGCAAGCTATTCCCTTGAAGCAATTA 1081
DB 4438 CTCTGTCCACTACCGGAGAGATCCCTTTTATGCGCAAGCTATTCCCTTGAAGTHATCA 4497
QY 1082 AGGGGGGGGACATCTCATCTTCTGCCACTCAAGGAAGTGCAGAGCTCGCGCAA 1141
DB 4498 AGGGGGGGGACATCTCATCTTCTGTCAATTCAAAGGAAGTGCAGAGAACTCGCGCAA 4557
QY 1142 AACTGTGCGGTGGCGGCTCAATGCGGTGCTTACTACCGGGGCTTGTGTGTCGTCA 1201
DB 4558 AGTGTGCTCATTTGGGCATCAATGCGGTGCTTACTACCGGGGCTTGTGTGTCGTCA 4617
QY 1202 TCCGACACAGTGTGAGTGTGCTGCGTGCAGTCAAGCCCTCATGACCGGCTTACCG 1261
DB 4618 TCCGACACAGCGGATGTTGCTGCGTGCAGTCAAGCCCTCATGACCGGCTTACCG 4677
QY 1262 GCGACTTCGATTCGGTGTAGACTGCAACACGCTGTGTACCCAGAGCTCGACTTCAGCC 1321
DB 4678 GCGACTTCGATTCGGTGTAGACTGCAATACGCTGTGTACCCAGAGCTCGACTTCAGCC 4737
QY 1322 TTGACCTTACCTTACCATTTAGACAAATCAGCTTCCCGAGTGTGTCTCCGTACTC 1381
DB 4738 TTGACCTTACCTTACCATTTAGACAAATCAGCTTCCCGAGTGTGTCTCCCGACTC 4797
QY 1382 AAGCTCGGGGTGAGACTGCGAGAGGAAGCAGGCACTACAGATTTGTGGCAGCGGG 1441
DB 4798 AAGCTCGGGGTGAGACTGCGAGAGGAAGCAGGCACTACAGATTTGTGGCAGCGGG 4857
QY 1442 AGCGTCTTCTTGGCATGTTTGAATCTGCTCTGCTGAGTGTGTATGACGGGGTGTG 1501
DB 4858 AGCGTCTTCTTGGCATGTTTGAATCTGCTCTGCTGAGTGTGTATGACGGGGTGTG 4917
QY 1502 CTTGGTATGAGCTTACCGCGCGGAGACACAGTGTAGGCTACGAGATACATGAACACCC 1561
DB 4918 CTTGGTATGAGCTTACCGCGCGGAGACTACAGTGTAGGCTACGAGGCTACATGAACACCC 4977
QY 1562 CGGGACTTCCCGTGTGCGAGACCATTTGAATTTTGGAGGGGCTTTTACGGGCTCA 1621
DB 4978 CGGGGCTTCCCGTGTGCGAGACCATTTGAATTTTGGAGGGGCTTTTACGGGCTCA 5037
QY 1622 CCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
DB 5038 CTATATAGTGCCTTCTTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC 5097
QY 1682 TGTAGCGTACCAAGCAGCGTGTGCTAGAGCTCAAGCCCTTCCCGTGTGGGAGCC 1741
DB 5098 TGTAGCGTACCAAGCAGCGTGTGCTAGAGCTCAAGCCCTTCCCGTGTGGGAGCC 5157
QY 1742 AGATGTGGAGTGTGTGATCCGCTTCAAGCCACCTTCAAGCCACCTTCAAGCCCTTCA 1801
DB 5158 AGATGTGGAGTGTGTGATCCGCTTCAAGCCACCTTCAAGCCACCTTCAAGCCCTTCA 5217

QY 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCCTGACGACCCAGTCAACAAGTATATCA 1861
DB 5218 ACAGACTGGGCGCTGTTCAGAAATGAATCAACCCTGACGACCCAGTCAACAATATCA 5277
QY 1862 TGACATGATGTCTGGCTGACCTGAGGTCGTCAAGTACCTGGTGTCTGTTGGCGG 1921
DB 5278 TGACATGATGTCTGGCGGACCTGAGGTCGTCAAGTACCTGGTGTCTGTTGGCGG 5337
QY 1922 TTCTGGCTGCTTTGGCGCGCTATTGCTTATCCACAGGCTGCTGCTCATAGTAGTAGGA 1981
DB 5338 TCCTGGCTGCTTTGGCGCGTATTGCTTCAACAGGCTGCTGCTCATAGTAGTAGGA 5397
QY 1982 TTGCTTGTTCGGAAGCGGCAATCATACCGGACGAGGAAGTCTCTTACCGGGAGTTG 2041
DB 5398 TCGTCTTGTTCGGAAGCGGCAATCATACCGGAGGTCCTCTTACCGAGAGTTG 5457
QY 2042 ATGAATGGAAGAGTGT 2059
DB 5458 ATGAGATGGAAGAGTGT 5475

RESULT 7

US-08-823-895A-25

; Sequence 25, Application US/08823895A

; Patent No. 6433159

; GENERAL INFORMATION:

; APPLICANT: Kevin P. Anderson

; TITLE OF INVENTION: Compositions And Methods For

; TITLE OF INVENTION: Treatment Of Hepatitis C Virus-Associated Diseases

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Jane Massey Licata, Esq.

; STREET: 66 E. Main Street

; CITY: Marlton

; STATE: NJ

; COUNTRY: USA

; ZIP: 08053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE

; COMPUTER: IBM 486

; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/823,895A

; FILING DATE: March 17, 1997

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/453,085

; FILING DATE: May 30, 1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/945,289

; FILING DATE: September 10, 1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Jane Massey Licata

; REGISTRATION NUMBER: 32,257

; REFERENCE/DOCKET NUMBER: ISPH-0203

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (609) 779-2400

; TELEFAX: (609) 810-1454

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9401

; TYPE: Nucleic

; STRANDEDNESS: Single

; TOPOLOGY: Linear

; ANTI-SENSE: NO

; US-08-823-895A-25

Query Match 86.7%; Score 1786; DB 3; Length 9401;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCCTATCATCGGCTATGCCAGAGGGGCGCTTTTGGGATGCATAATCACCA 61

Db 3418 TGGCGCCATCAACGGGTACGCCAGCAGCAAGGGGCTCTCAGGTGCAATAACCA 3477
 Qy 62 GCTTGACCGCGCGGAGCAAAACAGGTGAGGTGAGGTTCAGATCGTGTCAACTGCTG 121
 Db 3478 GCTTAACCGCGCGGAGCAAAACAGGTGAGGTGAGGTTCAGATCGTGTCAACTGCTG 3537
 Qy 122 CCAGACTTCTTGGCAACTGCATTAACGGGGTGTGTTGGAAGTGTCTACCAATGAGCG 181
 Db 3538 CCCAAACCTTCTGGCAACGTGCATCAATGGGGTGTGCTGGAAGTGTCTACCAAGGGCG 3597
 Qy 182 GAACAAGACCAATGCGTCACCTAAGGGTCTGTTCAGATGATACCAATGAGACC 241
 Db 3598 GAACGAGGACCAATGCGTCACCAAGGGTCTGTTCAGATGATACCAATGAGACC 3657
 Qy 242 AAGACCTGATAGCTGGCGGCTCCCAAGGTGCGGCTCATTAACCAATGAGCTGG 301
 Db 3658 AAGACCTTGTGGGCTGGCGGCTCCCGCAAGGTGAGCGCTCATTAACCAATGAGCTGG 3717
 Qy 302 GCTCTCGGACCTTACCTGTGTCAGAGGACGCGATGTCTCTGTGGCGCGGAGCG 361
 Db 3718 GCTCTCGGACCTTACCTGTGTCAGAGGACGCGATGTCTCTGTGGCGCGGAGCG 3777
 Qy 362 GTGATGCGAGGCGACCTGCTTTCCGCCCGGCTATCTCTTACTTGAAGGCTCTCTCG 421
 Db 3778 GTGATGCGAGGCGACCTGCTTTCCGCCCGGCTATCTCTTACTTGAAGGCTCTCTCG 3837
 Qy 422 GAGGCGCTCTGTGTCGCCCGCAGGACATGCGTAGGATATTCAGAGCGCGGTATGCA 481
 Db 3838 GGGGTCCGCTGTGTGTCGCCCGGCGGACGCGGTGGCAATATTTAGGCGCGGCTGTGCA 3897
 Qy 482 CCCGTGGAGTGTCTAAGGGGTGGACTTCATCCCGTAGAGAGCTTAGAGACCAACATGA 541
 Db 3898 CCCGTGGAGTGTCTAAGGGGTGGACTTCATCCCGTAGAGAGCTTAGAGACCAACATGA 3957
 Qy 542 GGTCCCGGTGTCTCAGCAAACTCTCTCCCAAGCAAGTGTGCGGAGCTTCCAAAGTGG 601
 Db 3958 GGTCCCGGTGTCTCAGCAAACTCTCTCCCAAGCAAGTGTGCGGAGCTTCCAAAGTGG 4017
 Qy 602 CCCACTGATGCTCCCAAGGCGAGGTGAAGACCAAGGTGCGGCGGATACGCGAG 661
 Db 4018 CTCACCTCCATGCTCCCAAGGCGAGGTGAAGACCAAGGTGCGGCGGATACGCGAG 4077
 Qy 662 CTCAGGGCTACAAGGTGCTGTGCTCAACCCCTCGTGTCTGCAAAATGGGCTTTGGTG 721
 Db 4078 CTCAGGGCTACAAGGTGCTGTGCTCAACCCCTCGTGTCTGCAAAATGGGCTTTGGTG 4137
 Qy 722 CTTCATGTCGAAGGCCAATGGATGATCTTAAATCAGGAGCTGGGGTGAGGACAAATTA 781
 Db 4138 CTTCATGTCGAAGGCCAATGGATGATCTTAAATCAGGAGCTGGGGTGAGGACAAATTA 4197
 Qy 782 CTACTGGGAGCGGATCAGTATTCACCTAGCGCAAGTTCCTTGGCGAGCGGGTGT 841
 Db 4198 CCACTGGGAGCGGATCAGTATTCACCTAGCGCAAGTTCCTTGGCGAGCGGGTGT 4257
 Qy 842 CAGGGGTGCTTTATGACATAAATTTGTGACAGTGGCCACTTCCAGGATGCAACATCA 901
 Db 4258 CAGGGGTGCTTTATGACATAAATTTGTGACAGTGGCCACTTCCAGGATGCAACATCA 4317
 Qy 902 TCTTGGGCATGGCACTGCTTGAACCAAGAGAGACCGCGGGGCGAGACTGACTGTGC 961
 Db 4318 TCTTGGGCATGGCACTGCTTGAACCAAGAGAGACTGCGGGGCGAGACTGCTTGTGC 4377
 Qy 962 TCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCGGCATCTTAACATCGAGGAGGTG 1021
 Db 4378 TCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCGGCATCTTAACATCGAGGAGGTG 4437
 Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAATTA 1081
 Db 4438 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAATTA 4497
 Qy 1082 AGGGGGGAGACATCTCATCTTCTGCGCACTCAAGAGAGTGGCGAGGCTCGCGCAA 1141

Db 4498 AGGGGGGAGACATCTCATCTTCTGTCTATTCAAAGAGAGTGCAGCAACTCGCGCAA 4557
 Qy 1142 AACTGGTTCGGTTCGGGCTCAATGCGGCTTACTACCGGCGCTTGAATGTCGGTCA 1201
 Db 4558 AGTGGTTCGGTTCGGGCTCAATGCGGCTTACTACCGGCGCTTGAATGTCGGTCA 4617
 Qy 1202 TCCCGACAGTGGTTCGGGCTTGTGTCGGCAACTGAGCGGCTCATGATGCGGCTTTACCG 1261
 Db 4618 TCCCGACAGTGGTTCGGGCTTGTGTCGGCAACTGAGCGGCTCATGATGCGGCTTTACCG 4677
 Qy 1262 GCGACTTCGATTCGGTTCGGTTCGGCAACTGAGCGGCTCATGATGCGGCTTTACCG 1321
 Db 4678 GCGACTTCGATTCGGTTCGGTTCGGCAACTGAGCGGCTCATGATGCGGCTTTACCG 4737
 Qy 1322 TTGACCTTACCTTCAGCAATGAGCAACTGAGCGGCTCATGATGCGGCTTTACCG 1381
 Db 4738 TTGACCTTACCTTCAGCAATGAGCAACTGAGCGGCTCATGATGCGGCTTTACCG 4797
 Qy 1382 AACGTTCGGGTAGGACTGGCAGAGGGAAGCAGGATCTTACAGATTTTGTGGCACCGGG 1441
 Db 4798 AACGTTCGGGTAGGACTGGCAGAGGGAAGCAGGATCTTACAGATTTTGTGGCACCGGG 4857
 Qy 1442 AGCGTCTTTCGGGATGTTGACTGCTGCTCTGCGAGTGTCTGAGCGGGTGTG 1501
 Db 4858 AGCGTCTTTCGGGATGTTGACTGCTGCTCTGCGAGTGTCTGAGCGGGTGTG 4917
 Qy 1502 CTTGTGATGAGTTCAGCGCCGCGGAGCAAGTTCAGATTTTGTGGAGGCGCTTTTACGGTCTCA 1561
 Db 4918 CTTGTGATGAGTTCAGCGCCGCGGAGCAAGTTCAGATTTTGTGGAGGCGCTTTTACGGTCTCA 4977
 Qy 1562 CCGGACTTTCGGGTTCAGCAAGCAACTTTCAGATTTTGTGGAGGCGCTTTTACGGTCTCA 1621
 Db 4978 CCGGACTTTCGGGTTCAGCAAGCAACTTTCAGATTTTGTGGAGGCGCTTTTACGGTCTCA 5037
 Qy 1622 CCCATAGAGCGGCTTTCCTTATCCAGCAAGAGAGTGGGGAACCTTTCCTTATC 1681
 Db 5038 CTCATATAGAGCGGCTTTCCTTATCCAGCAAGAGAGTGGGGAACCTTTCCTTATC 5097
 Qy 1682 TGGTAGCTTCAGAGCGGCTTTCCTTATCCAGCAAGAGAGTGGGGAACCTTTCCTTATC 1741
 Db 5098 TGGTAGCTTCAGAGCGGCTTTCCTTATCCAGCAAGAGAGTGGGGAACCTTTCCTTATC 5157
 Qy 1742 AGATGGAAGTTCCTTATCCAGCGGCTTTCCTTATCCAGCGGCAACCTTTCCTTATC 1801
 Db 5158 AGATGGAAGTTCCTTATCCAGCGGCTTTCCTTATCCAGCGGCAACCTTTCCTTATC 5217
 Qy 1802 ATAGCTGGGCTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 1861
 Db 5218 ACAGCTGGGCTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 5277
 Qy 1862 TGACATGATTCGGGCTGACCTGGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 1921
 Db 5278 TGACATGATTCGGGCTGACCTGGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 5337
 Qy 1922 TTCTGGCTTTCGGGCTGATTCCTTATCCAGCGGCTTCAGAGTTCAGAGTTCAGAGTTCAGAG 1981
 Db 5338 TTCTGGCTTTCGGGCTGATTCCTTATCCAGCGGCTTCAGAGTTCAGAGTTCAGAGTTCAGAG 5397
 Qy 1982 TTGCTTTCGGGAAAGCGGCAATTCATCCGAGAGGAGTTCCTTACCGGGAGTTCG 2041
 Db 5398 TCGTCTTCGGGAAAGCGGCAATTCATCCGAGAGGAGTTCCTTACCGGGAGTTCG 5457
 Qy 2042 ATGAATGGAAGTTCG 2059
 Db 5458 ATGAATGGAAGTTCG 5475

RESULT 8
 PCT-US91-02225-9
 ; Sequence 9, Application PC/TUS9102225
 ; GENERAL INFORMATION:
 ; APPLICANT: HOUGHTON, MICHAEL
 ; APPLICANT: CHOO, QUI-LIM

APPLICANT: KUO, GEORGE

TITLE OF INVENTION: COMBINATIONS OF HEPATITIS C VIRUS

TITLE OF INVENTION: COMBINATIONS OF HEPATITIS C VIRUS

TITLE OF INVENTION:

TITLE OF INVENTION: ANT

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison & Foerster

STREET: 545 Midd.

CITY: Merlo

STATE: CA

COUNTRY: US

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Polycase #1.0 Version #1.25

SOFTWARE: PATENT IN REL
CURRENT APPLICATION DATA:

CURRENT AFFILIATION DATA:
APPLICATION NUMBER: PCT/US91/02225

AFFILIATION NUMBER: F
FILING DATE: 19910329

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CIOTTI, THOMAS E.

REGISTRATION NUMBER: 21,013

REFERENCE/DOCKET NUMBER: 2300-0101.44

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 327-7255

TELEFAX: (415) 327-2951

TELEX: 706141

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 9401 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

JS91-02225-9

Query Match	86.6%	Score 1785.2	DB 5	Length 9401
Best Local Similarity	91.5%	Pred. No. 0		
Matches 1883	Conservative	6	Mismatches 169	Indels 0
Gaps	0			
QY	2	TGGCGGCTATCAGCGGCTATGCCAGCAGACAGGGGCCCTTTTGGGATGCATTAATCACCA	61	
DB	3418	TGGCGGCCATCAGCGGCTAGCCGACGACAGAGGGGCCCTCTAGGTGTCATAATCACCA	3477	
QY	62	GCTTGAACCGCCGGGACAAAAACAGTGCAGGGTGAGGTTTCAGATCGTGTCAACTGCTG	121	
DB	3478	GCTTAATGCGCCGGGACAAAAACAGTGAGGGTGAGGTCAGATGTTGTCAACTGCTG	3537	
QY	122	CCCAGACTTTCCTTGGCAACCTGCATTAACGGGGTGTTGAGACTGTCTACATCGGAGCCG	181	
DB	3538	CCCAACCTTCTGGGCACACGTGCATCAATGGGGTGCTGACTGTCTACACGGGGCCG	3597	
QY	182	GAACAAAGGACCATTCGGTCACTAAGGGTCCTGTTATCCAGATGTACACCAATGTGGACC	241	
DB	3598	GAACGAGGACCATCGGTCACCAAGGGTCCTGTTCATCCAGATGTATACCAATGTAGACC	3657	
QY	242	AAGACCTGTAGAGCTGGCCGCTCCCAAGGTGCCGCTCATTTAAACCAATGCATTTGCG	301	
DB	3658	AAGACCTGTGTGGCTGGCCGCTTCGCAAGGTATSCCGCTCATTTGACACCCCTGCATTTGCG	3717	
QY	302	GCTTCCTCGACCTTTTACTCGTTCACGAGGCACGCCGATGTCATTTCCTGTGCGCCGACGG	361	
DB	3718	GCTTCCTCGACCTTTTACTCGTTCACGAGGCACGCCGATGTCATTTCCTGTGCGCCGCGGG	3777	
QY	362	GTGATGGCAGGGGCAGCCTCTTTTCGCCCGGCCCTATCTCTTACTTCAAAAGGCTCTCTCGG	421	
DB	3778	GTGATAGCAGGGGCAGCCTCTGTGTCGCCCGGCCCATTTCTACTTGAAGGGCTCCTCGG	3837	
QY	422	GAGGCCCTCTGTGTGCCCCCGCAGGACATGCCGTAGGCATATTACAGAGCCGCGGTATGCA	481	
DB	3838	GGGGTCCGCTGTTGTGCCCGCGGGGCACGCCGTGGGCATATTTAGGGCCGCGGTGCA	3897	

QY 1562 CGGGACTTCCGCTGTCAGACCATCTTGAATTTGGAGGGCGCTTTAGGGCTCTCA 1621
Db 4978 CGGGGCTTCCGCTGTCAGACCATCTTGAATTTGGAGGGCGCTTTAGGGCTCTCA 5037
QY 1622 CCCACATAGACGCCACCTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
Db 5038 CTCATATAGATGCCACCTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCTTACC 5097
QY 1682 TGGTAGCGTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTGTGGGACC 1741
Db 5098 TGGTAGCGTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTGTGGGACC 5157
QY 1742 AGATGTGGAAGTGTGATCGCTCTCAAGCCCAACCTCCATGGGCAACACCTCTGTAT 1801
Db 5158 AGATGTGGAAGTGTGATCGCTCTCAAGCCCAACCTCCATGGGCAACACCTCTGTAT 5217
QY 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAAGCTCAAGCCCAACCTCCAGTATATCA 1861
Db 5218 ACAGACTGGGCGCTGTCCAGAAATGAAGTCAAGCTCAAGCCCAACCTCCAGTATATCA 5277
QY 1862 TGACATGTATGTGGCTGACCTGGAGTGTCTACAGTACCTGGGTGCTCTGTGGGGCG 1921
Db 5278 TGACATGTATGTGGGCGGACCTGGAGTGTCTACAGTACCTGGGTGCTCTGTGGGGCG 5337
QY 1922 TTCTGGCTGCTTTGGGCGGCTATTGCTATCAAGGCTGCTGTATAGTAGTAGGA 1981
Db 5338 TCTGTGCTGCTTTGGGCGGCTATTGCTGTCAAGGCTGCTGTATAGTAGTAGGA 5397
QY 1982 TTGTCTTGTCCGAAAGCCGCAATCATACCGACAGGAAAGTCTCTTACCGGGAGTTCG 2041
Db 5398 TCGTCTTGTCCGAAAGCCGCAATCATACCTGACAGGAAAGTCTCTTACCGGAGTTCG 5457
QY 2042 ATGAATGAAGAGTCT 2059
Db 5458 ATGAGATGAAGAGTCT 5475

RESULT 9

US-08-444-818-65
; Sequence 65, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 011002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 6785 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3...6785
; US-08-444-818-65

Query Match 86.6%; Score 1784.4; DB 3; Length 6785;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 2 TGGCGCTATACACGGCTATGCCAGCAGACAAGGGGCTTTTGGATGATATACCA 61
Db 1204 TGGCGCCCATACCGCGCTACGCCAGCAGACAAGGGGCTTCTAGGTCATATACCA 1263
QY 62 GCTTGAACGGCGCGGACAAACCAAGGTGGAGGTTCAGATCGTGTCACTGCTG 121
Db 1264 GCCTAACTGGCGCGGACAAACCAAGTGGAGGTTCAGATCGTGTCACTGCTG 1323
QY 122 CCCAGACTTTCTTGGCAACCTGCAATTAACGGGTGTGTGGACTGTCTACCATGAGCG 181
Db 1324 CCCAAACCTTCTTGGCAACCTGCAATTAACGGGTGTGTGGACTGTCTACCATGAGCG 1383
QY 182 GAACAAGGACCATTTGCGTACCTTAAGGTCTCTGTATCCAGATGTACACCAATGTGACC 241
Db 1384 GAACAGGACCATTCGCTGCTACCAAGGTCTCTGTATCCAGATGTATACCAATGTAGACC 1443
QY 242 AAGACCTCGTAGGCTGCGCCGCTCCCAAGGTGCGGCTCATTAACCACTGCACTTGGC 301
Db 1444 AAGACCTTGTGGGTGCGCCGCTCCGCAAGTAGCGCTCATTTGACACCTTGCCTTGGC 1503
QY 302 GCTCTCGGACTTTACTCTGTGACAGGACGCGGATGTCATTCCTGTGCGCGACGCGG 361
Db 1504 GCTCTCGGACTTTACTCTGTGACAGGACGCGGATGTCATTCCTGTGCGCGCGCGG 1563
QY 362 GTGTGGCAGGCGAGCTCTTTCGCCCCCGGCTATCTTCTTACTTGAAGGCTCTCTCG 421
Db 1564 GTGTAGCAGGCGGACCTGCTGTGCGCCCCGCGCCATTTCTTACTTGAAGGCTCTCTCG 1623
QY 422 GAGGCTCTGCTGTGCGCGGACGACATGCGGTAGGATATTAGAGCCCGGGTATGCA 481
Db 1624 GGGGTCCGCTGTGTGCGCGCGGACGCGGCTATTTAGGCGCGCGGTGTGCA 1683
QY 482 CCCGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCGTAGAGAGCTTAGAGACAACTGA 541
Db 1684 CCCGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCGTAGAGAGCTTAGAGACAACTGA 1743
QY 542 GGTCCCCCGGTGTTCTCAGACAACTCTCTCCCAAGCAGTGCCTCCAGAGCTTACCAAGTGG 601
Db 1744 GGTCCCCCGGTGTTACCGGATAACTCTCTCCACAGTAGTAGTCCCGCAGAGCTTCCAGGTGG 1803
QY 602 CCCACTGTGATCTCCACCGGACGCGTAGAGACCAAGTCCCGCGCGCATACGAG 661
Db 1804 CTCACCTCCATGCTCCACAGCAGCGGCAAAAGCAGGTCCTCCGCTGCTATGAG 1863
QY 662 CTCAGGGCTTACAAGGTGTGCTGTCTCAACCCCTCCGTTGCTGCAACAATTTGGCTTGGTG 721
Db 1864 CTCAGGGCTTACAAGGTGTGCTGTCTCAACCCCTCTGTTGCTGCAACAATTTGGCTTGGTG 1923
QY 722 CTTACATGTCCAAGGCCCATGGATGTGATCTTAAATCATGAGGCTGGGTGAGGACAATTA 781
Db 1924 CTTACATGTCCAAGGCTCATGGATGCTTAAATCATGAGGACCGGGGTGAGACAATTA 1983
QY 782 CTACTGGGACCGGATCAGTATTCCACTAGCGCAAGTTCCTTCCGACGCGGGTGT 841
Db 1984 CCACTGGCAGCCCCATCAGTACTCAACCTTCTGCTGCAACAATTTCTTCCGACGCGGGTGT 2043
QY 842 CAGGGGGTCTTTATGACATAATAATTTGTGAGAGTGGCCACTTCCACGGATGCAATCA 901

2044	CGGGGGCGCTTATGACATAAATAAATTTGTGACAGCTGCCACTCCACGGATGCCACATCCA	2101
QY	TC TTGGGCATTTGGCACTGTCTCTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGTC	961
DB	TC TTGGGCATCGGCACCTGTCTTGACCAAGCAGAGACTGCGGGGGGAGACTGGTTGTGC	2163
QY	TCdCCACCGCTACCCCTCCGGGCTCGTCTACTGTGCCCAATCTTAACATCGAGGAGGTTG	1021
DB	TCdCCACCGCACCCCTCCGGGCTCGTCTACTGTGCCCAATCCCAACATCGAGGAGGTTG	2223
QY	CTGTGTCCACTACCGGAGAGATCCCCTTTTATGCGAAGGCTATTTCCTCTGAAGCAATTA	1081
DB	CTGTGTCCACACCGGAGAGATCCCCTTTTACGCAAGGCTATCCCTCTGGAATATCA	2283
QY	AGdGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAGAGTGCAGAGCTCGCGCAA	1141
DB	AGGGGGGAGACATCTCATCTTCTGTCATTCAAAGAGAGTGCAGAGCTCGCGCAA	2343
QY	AAGTGTGCGGTTGGGGCTCAATGCGGTGCTTACTACCGGGGCTTGATGTGTCGTCA	1201
DB	AGTGTGTGCAATTTGGGCATCAATGCCGTGCGCTACTACCGGGCTTGTGACGTGCGCTCA	2403
QY	TCdCGACACTGGTGTGCTGCTGCTGCGCAACTGACGCGCTCATGACCGGCTTTACCG	1261
DB	TCdCGACACGGCGGATGTTGTGCTGCTGCGCAACCGATGCCCTCATGACCGGCTATACCG	2463
QY	GdJACTTCGATTCGGGTATAGACTGCAACACGTGTGTACCCAGACAGTGCAGCTTCAGCC	1321
DB	GdJACTTCGACTCGGTGATAGACTGCANATGCTGTACCCAGACAGTGCAGCTTCAGCC	2523
QY	TTdACCCTACTTTCACCATTTGAGACATACGCTTCCCAAGATGCTCTCCCGACTC	1381
DB	TTdACCCTACTTTCACCATTTGAGACATACGCTTCCCAAGATGCTCTCCCGACTC	2583
QY	AAcGTTCGGGTAGGACTGGCAGAGGAGCCAGGCATCTACAGATTTGTGGCACCGGGG	1441
DB	AAcGTTCGGGACGAGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGGG	2643
QY	AGCGTCTCTTCGCACTGTTTGACTTCGTCGTCTCTGCGAGTGTATGACGCGGGTTGTG	1501
DB	AGCGCTCTTCGCGATGTTTCGACTTCGTCGTCTCTGCGAGTGTATGACGCGGGTTGTG	2703
QY	CTTGGTATGAGCTTACGCCCGCGAGACCAcAGTTAGGCTACGAGCATACATGAACACC	1561
DB	CTTGGTATGAGCTACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC	2763
QY	CGGACTTCCTCGGTGTGCAAGCCATCTTGAAATTTTGGGAGGGGCTTTTACCGGCTCA	1621
DB	CGGGGCTTCCTCGGTGTGCGAGACCATCTTGAAATTTTGGGAGGGGCTTTTACAGGGCTCA	2823
QY	CCACATAGACGCCACTTCTCTATCCCAGACAAGACAGAGTGGGGAACCTTTCCCTATC	1681
DB	CTCATATAGATGCCACTTCTATATCCAGACAAGACAGAGTGGGGAACCTTTCTTACC	2883
QY	TGdTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCTCCCGCTCGTGGGACC	1741
DB	TGGTAGCGTACCAAGCCACCGTGTGCGCTAGGCTCAAGCCCTCTCCCGCTCGTGGGACC	2943
QY	AGATGTGGAAAGTGTGATCCGCTCTCAAGCCCACTCTGATGGGCCCAACCTCTGCTAT	1801
DB	AGATGTGGAAAGTGTGATTTGATTTGCGCTCAAGCCCACTCTGATGGGCCCAACCTCTGCTAT	3003
QY	ATAGACTGGGCGTGTCCAGAAATGAGTCACTGACGCACCCAGTCAACCAAGTATATCA	1861
DB	ACAGACTGGGCGTGTTCAGAAATGAAATACCTCTGACGCACCCAGTCAACCAATATATCA	3063
QY	TGACATGATGTCCGCTGACCTGGAGGTTCGTCAcAGTACCTGGGTCTCGTTGGGGCG	1921
DB	TGACATGATGTCCGCGCACTGGAGGTTCGTCAcAGCACTGGGTCTCGTTGGGGCG	3123
QY	TTCTGGCTGTCTTGGCGCGCTATTTGCTATCCAAGGCTCGTGGTCAATAGTAGGTAGGA	1981
DB	TCCTGGCTGTCTTGGCGCGCTATTTGCTGTCAACAGGCTCGTGGTCAATAGTGGGAGGG	3183

Qy	1982	TTGTCTTGTCCGGAAGCCGCGCAATCATATCCCGACAGGGAGTCTCTACCGGAGTTGC	2041
Db	3184	TCGTCTTGTTCGCGGAAGCCGCGCAATCATATCTCACAGGGAAGTCTCTACCGAGAGTTGC	3243
Qy	2042	ATGAAATGGAAGAGTGTCT	2059
Db	3244	ATGAGATGGAAGAGTGTCT	3261

```

RESULT 10
US-08-444-818-88
; Sequence 88, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8316 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8316
US-08-444-818-88

```

	Query Match	86.6%	Score 1784.4	DB 3	Length 8316
	Best Local Similarity	91.7%	Pred. No. 0		
	Matches 1887	Conservative	0	Mismatches 171	Indels 0
	Gaps	0			
Qy	2	TGGCGCCTATCATCGGCCCTATGCCAGCAGACAAGGGGCCCTTTTGGGATGCATATACCA	61		
Db	2735	TGGCGCCCATCATCGGCGTACGCCAGCAGACAAGGGGCCCTCTTAGGTTGCATATACCA	2794		
Qy	62	GCTTGAACCGCCCGGGACAAAAACAGGTGGAGGGTGAGGTTTCAGATCGTGTCAACTGCTG	121		
Db	2795	GCCTAACTGCGCGGACAAAAACCAAGTGAGGGTGAGGTCAGATTTGTCAACTGCTG	2854		
Qy	122	CCGAGCTTCTTGGGCAACCTGCATTAACCGGGGTGTGTTGGACTGTCTACCATGGAGCCG	181		
Db	2855	CCCAACCTTCTTGGCAACGTGATCAATGGGGTGTGTGGACTGTCTACACGGGGCCG	2914		

QY	182	GAACAAGNACATTTCGGTCACCTAAGGGTCCTGTTATCCAGATGTACACCAATGTGGACC	241
Db	2915	GAACGAGGACCAATCGCGTCACCAAGGGTCCTGTATCCAGATGTATACCAATGTAGACC	2974
QY	242	AAGACCTCGTAGCGCTGGCCCGCTCCCAAGGTGCCCGTCAATTAACCATGCACTTCGG	301
Db	2975	AAGACCTTGTGGCTGGCCCGCTCCGCAAGGTAGCGCTCATTTGACACCTGCACTTCGG	3034
QY	302	GTCCTCGGACCTTTACTCTGGTCAAGAGCAGCCGATGTATTTCTGTGGCCGCAACGG	361
Db	3035	GCTCTCGGACCTTTACTCTGGTCAAGAGCAGCCGATGTATTTCTGTGGCCGCGCGGG	3094
QY	362	GTGATGGCAGGGCAGCTCTTTCCGCCCGCCTATCTCTACTTTGAAGGCTCCTCGG	421
Db	3095	GTGATAGCAGGGCAGCTCTGTTCGCCCGGCCCATTTCTCTATTGAAGGCTCCTCGG	3154
QY	422	GAGGCCCTCTGTGTGCCCGCAGGACATGCCGTAGGCATATTCAGAGCCGCGGTATGCA	481
Db	3155	GGGTCCTGTGTGTGCCCGCGGGCAGCCGTGGGCATATTTAGGGCCGCGGTGTGCA	3214
QY	482	CCCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGTAGAGAGCTTAGAGACAACCATGA	541
Db	3215	CCCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATGA	3274
QY	542	GGTCCCGGTGTTCTCAGACAACCTCTCCCCACAGCAGTGCCTCCAGAGCTACCAAGTGG	601
Db	3275	GGTCCCGGTGTTCTACGGATACTCTCTCCACCACTAGTGTCCCCAGAGCTTCAGGTGG	3334
QY	602	CCCACCTGCATGCTCCACCGGCAGCGGTAAGAGCACCAAGGTCGCGGCGCATGCGAG	661
Db	3335	CTCACCTCCATGCTCCACAGGCAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCAG	3394
QY	662	CTCAGGGCTACAAGGTGTGGTGCTCAACCCCTTCGTTGTGTCGCAACAATGGGCTTTGG	721
Db	3395	CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGTCGCAACACTGGGCTTTGGTG	3454
QY	722	CTTACATGTCGAAGGCCCATGGATTCATCTTAACATCAGGATGCGGGTGAGACAATTA	781
Db	3455	CTTACATGTCGAAGGCTCATGGATCGATCTTAACATCAGGACCGGGTGAGACAATTA	3514
QY	782	CTACTGSCAGCCCGATCACGTAATCCACCTACGGCAAGTTCCTTGGCGACGGCGGGTGT	841
Db	3515	CCACTGSCAGCCCGATCACGTAATCCACCTACGGCAAGTTCCTTGGCGACGGCGGGTGT	3574
QY	842	CAGGGGTGTTATGACATAATTAATTTGTGACAGGTGCCACTCCAGGATGCAACATCCA	901
Db	3575	CAGGGGGCGGTATGACATAATTAATTTGTGACAGGTGCCACTCCAGGATGCCACATCCA	3634
QY	902	TCATTGGGCATTTGGCACTGTCTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC	961
Db	3635	TCATTGGGCATTCGGCACTGTCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTGC	3694
QY	962	TCGCCACCGCTPACCCCTCCGGGCTCCGTCACTGTGCCCATTCCTTAACATCGAGAGGTTG	1021
Db	3695	TCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATTCGCCCAACATCGAGAGGTTG	3754
QY	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTTCCTCTTGAAGCAATTA	1081
Db	3755	CTCTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTGAAGTAATCA	3814
QY	1082	AGGGGGGAGACATCTCATCTTCGCCACTCAAAAGAAAGTAGCGAGCTCGCGCAA	1141
Db	3815	AGGGGGGAGACATCTCATCTTCGTTCATTCAAAGAAAGTAGCGAGCTCGCGCAA	3874
QY	1142	AACGTGTCGGCTTGGGGCTCAATCGCGTGGCTTACTACGGCGGCTTGATGTGTCGCTCA	1201
Db	3875	AGCTGGTCGCATTTGGCATCAATGCCGTGGGCTACTACCGCGGTCTTGAAGTGTGTCGCTCA	3934
QY	1202	TCCCGACCAAGTGTGACGTTCTGCTGTGGCAACTGACGCCCTCATGACCGGGCTTTACCG	1261
Db	3935	TCCCGACCAAGCGGCGATGTGTGCTGTGGCAACGATGCGCTCATGACCGGCTATACCG	3994
QY	1262	CGCACTTCGATTCGGTGTATAGACTGCACAACGTTGTCAACCGACAGTGTGATTCAGCC	1321

Db	3995	GCACCTTCGACTCGGTGTAGACTGCAATACGCTGTGTACCCAGACAGTCGATTTTCAGCC	4054
Qy	1322	TTGACCCCTACCTTCACATTTGAGACAAATCAGCTTCCCAGGATGCTGTCTCCCGTACTTC	1381
Db	4055	TTGACCCCTACCTTCACATTTGAGACAAATCAGCTTCCCAGGATGCTGTCTCCCGACATC	4114
Qy	1382	AACGTTCGGGTAGGACTTGGCAGAGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGGG	1441
Db	4115	AACTTCGGGCGCAGACTTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGGG	4174
Qy	1442	AGCGTCCTTCTGCGCATGTTTGAATCTGTCGTCTCTGCGAGTGTCTATGACGCGGGTTGTG	1501
Db	4175	AGCGCCCTCCGCGCATGTTTGAATCTGTCGTCTCTGCGAGTGTCTATGACGCGGGTTGTG	4234
Qy	1502	CTTTGGTATGACTTACGCGCGCGAGACCAACAGTAGGCTACGAGCATACATGAACACCC	1561
Db	4235	CTTTGGTATGACTTACGCGCGCGAGACTACAGTAGGCTACGAGCATACATGAACACCC	4294
Qy	1562	CGGGACTTCCTCGTGTGCCAAGACCATCTTGAAATTTTGGGAGGGCGCTTTTACGGGTCTCA	1621
Db	4295	CGGGGCTTCCTCGTGTGCCAGGACCATCTTGAAATTTTGGGAGGGCGCTTTTACAGGCTCA	4354
Qy	1622	CCCACTAGACGCCCACTTCTATCCACAGACAAAGCAGAGTGGGGAAAACCTTCCCTATC	1681
Db	4355	CTCATATAGATGCCCACTTCTATCCACAGACAAAGCAGAGTGGGGAGAACCCTTCTTACC	4414
Qy	1682	TGTTAGGTTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCTCCCGCTCGTGGGACC	1741
Db	4415	TGTTAGGTTACCAAGCACCGTGTGGCTTAGGGCTCAAGCCCTCTCCCGCTCGTGGGACC	4474
Qy	1742	AGATGTGGAAGTGTTCATCCGCTCTCAAGCCCACTTCCATGGGCCCAACACCTCTGCTAT	1801
Db	4475	AGATGTGGAAGTGTTCATTCGCTCTCAAGCCCACTTCCATGGGCCCAACACCTCTGCTAT	4534
Qy	1802	ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGCAACCAAGTCAACCAATATCA	1861
Db	4535	ACAGACTGGGCGCTGTTCAGAAATGAATCAACCTGACGCAACCAAGTCAACCAATATCA	4594
Qy	1862	TGACATGTATGTCCGGCTGACCTGGAGTCTGTCAAGTACCTGGGTCTCTGTTGGCGGCG	1921
Db	4595	TGACATGTATGTCCGGCGACCTGGAGTCTGTCAAGTACCTGGGTCTCTGTTGGCGGCG	4654
Qy	1922	TTCTGGCTGTCTTGGCGCGCTATTGCTTATCCAGAGCTCGTGGTCAATAGTAGGTAGGA	1981
Db	4655	TCCTGGCTGTCTTGGCGCGCTATTGCTTATCCAGAGCTCGTGGTCAATAGTAGGTAGGA	4714
Qy	1982	TTGTCTTGTCCGGAAAGCCGCAATCATACCCGACAGGGAAGTCTCTTACCGGGAGTTGG	2041
Db	4715	TCGTCTTGTCCGGGAAGCCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTGG	4774
Qy	2042	ATGAAATGGAAGAGTGCT	2059
Db	4775	ATGAGATGGAAGAGTGCT	4792

```

RESULT 11
US-08-444-818-137
; Sequence 137, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:

```


;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/444,818
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/403,590
;; FILING DATE: 14-MAR-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Harbin, Alisa A.
;; REGISTRATION NUMBER: 33,895
;; REFERENCE/DOCKET NUMBER: 0110.002
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (508)359-3876
;; TELEFAX: (508)359-3885
;; INFORMATION FOR SEQ ID NO: 137:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 8987 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..8985
US-08-444-818-137

Query Match 86.6%; Score 1784.4; DB 3; Length 8987;

Best Local Similarity 91.7%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY	2	TGGGGCTATCAGGGCTATGCCAGCAGACAGGGGCTTTGGGATGCATTAATCACCA	61
DB	3077	TGGGGCTATCAGGGCTATGCCAGCAGACAGGGGCTTTGGGATGCATTAATCACCA	3136
QY	62	GCTTTGACCGCGCGGACAAAACAGGTGAGGGTTCAGATCGTGTCAACTGCTG	121
DB	3137	GCCTAACTGCGCGGACAAAACAGGTGAGGGTTCAGATCGTGTCAACTGCTG	3196
QY	122	CCAGACTTTCTTGGCAACTGCAATTAACGGGGTGTGTTGGAATGTCTACATGAGCGG	181
DB	3197	CCCAAACTTTCTTGGCAACTGCAATTAACGGGGTGTGTTGGAATGTCTACATGAGCGG	3256
QY	182	GAACAGGACCATTTGGTTCACCTTAAGGTTCTGTATCCAGATGTACCAATGTGGAC	241
DB	3257	GAACAGGACCATTTGGTTCACCTTAAGGTTCTGTATCCAGATGTATACCAATGTAGAC	3316
QY	242	AAGACCTCGTAGCTGGCGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGGC	301
DB	3317	AAGACCTTGTGGCTGGCGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGGC	3376
QY	302	GCTCTCGGACCTTTACCTGGTCAAGGACACCGCGATGTCATCTCTGTGCGCGGCGG	361
DB	3377	GCTCTCGGACCTTTACCTGGTCAAGGACACCGCGATGTCATCTCTGTGCGCGGCGG	3436
QY	362	GTATGCGGAGGCGAGCTGCTTTCGCCCGGCTATCTTACTTTGAAGGCTCTCTCG	421
DB	3437	GTATGCGGAGGCGAGCTGCTTTCGCCCGGCTATCTTACTTTGAAGGCTCTCTCG	3496
QY	422	GAGGCTCTGTGTGCGCGGACAGCATGCGGTAGGCATATTACAGAGCGCGGTATGCA	481
DB	3497	GAGGCTCTGTGTGCGCGGACAGCATGCGGTAGGCATATTACAGAGCGCGGTATGCA	3556
QY	482	CCCGTGGAGTGGCTTAAGCGGCTGGACTTTCATCTCCCGTGGAGCTTTAGAGACAACATGA	541
DB	3557	CCCGTGGAGTGGCTTAAGCGGCTGGACTTTCATCTCCCGTGGAGCTTTAGAGACAACATGA	3616
QY	542	GGTCCCCGGTGTCTCAGACAATCTCTCCCAACAGAGTGCCTCCAGAGTCAACAGTGG	601
DB	3617	GGTCCCCGGTGTCTCAGACAATCTCTCTCCCAACAGAGTGCCTCCAGAGTCAACAGTGG	3676

QY	602	CCCACCTGCATGCTCCACCGCAGCGGTAAAGAGCACCAAGGTCCCGGCGCATACGCAG	661
DB	3677	CTCACCTCCATGCTCCACCGCAGCGGTAAAGAGCACCAAGGTCCCGGCGCATATGCAG	3736
QY	662	CTCAGGGCTACAAGGTGCTGCTCAACCCCTCCGTGCTGCAACAATGGGGTTGGTG	721
DB	3737	CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACAATGGGGTTGGTG	3796
QY	722	CTTACATGTCGAAGGCCATGGATTGATCTCTAAATCAGGACTGGGGTGAGACAATTA	781
DB	3797	CTTACATGTCGAAGGCCATGGATTGATCTCTAAATCAGGACTGGGGTGAGACAATTA	3856
QY	782	CTACTGGCAGCCCGATCACGTATTCCACCTACGGCAAGTTCTTTCGCGAGCGGGTGT	841
DB	3857	CCACTGGCAGCCCGATCACGTATTCCACCTACGGCAAGTTCTTTCGCGAGCGGGTGT	3916
QY	842	CAGGGGTGCTTTATGACATAAATTTGTGACAGGTGCCACTCCACGGATGCAACATCCA	901
DB	3917	CGGGGGGCGCTTTATGACATAAATTTGTGACAGGTGCCACTCCACGGATGCAACATCCA	3976
QY	902	TCTTGGGCATTTGGCACTGCTTACCAAGCAGAGACCGGGGGGAGACTCACCTGTGC	961
DB	3977	TCTTGGGCATTTGGCACTGCTTACCAAGCAGAGACTGGGGGGGAGACTCGTGTGC	4036
QY	962	TCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCTCTAAATCAGAGAGGTG	1021
DB	4037	TCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCTCTAAATCAGAGAGGTG	4096
QY	1022	CTCTGTCTCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCCTTTGAAGCAATTA	1081
DB	4097	CTCTGTCTCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCCTTTGAAGCAATTA	4156
QY	1082	AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAAAGTGCAGCAGCTCGCGCAA	1141
DB	4157	AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAAAGTGCAGCAGCTCGCGCAA	4216
QY	1142	AACTGTCTCGCTTGGCGCTCAATTCGCGTGGCTTACTACCGCGGCTTGTGTCCGTCA	1201
DB	4217	AGCTGTCTCGCTTGGCGCTCAATTCGCGTGGCTTACTACCGCGGCTTGTGTCCGTCA	4276
QY	1202	TCCGACAGAGTGTGAGCTGTGCTGTGGCAATCGACGCCCTCATGACCGGTTTACCG	1261
DB	4277	TCCGACAGAGTGTGAGCTGTGCTGTGGCAATCGATGCCCTCATGACCGGTTTACCG	4336
QY	1262	GGGACTTTCGCTTGGGTAGACTGCAACACGCTGTGTCAACCGACAGCTCGACTTCAGCC	1321
DB	4337	GGGACTTTCGCTTGGGTAGACTGCAACACGCTGTGTCAACCGACAGCTCGACTTCAGCC	4396
QY	1322	TTGACCTTACCTTTCACCATTTAGACAATCACCGCTTCCCGAGGATGCTCTCCCGTACTC	1381
DB	4397	TTGACCTTACCTTTCACCATTTAGACAATCACCGCTTCCCGAGGATGCTCTCCCGTACTC	4456
QY	1382	AACGTCGGGGTAGGACTGGCAGAGGGAAGCAGGCAATCTACAGATTTGTGGCAACCGGGG	1441
DB	4457	AACGTCGGGGTAGGACTGGCAGAGGGAAGCAGGCAATCTACAGATTTGTGGCAACCGGGG	4516
QY	1442	AGCGCTCTTCTGCAATGTTTGTGCTGCTCTGCGAGTGTATGACGCGGGTGTG	1501
DB	4517	AGCGCTCTTCTGCAATGTTTGTGCTGCTCTGCGAGTGTATGACGCGGGTGTG	4576
QY	1502	CTTGGTATGAGCTTACCGCGCGGAGACCAACATAGGCTACGAGATACATGAACACCC	1561
DB	4577	CTTGGTATGAGCTTACCGCGCGGAGACCAACATAGGCTACGAGATACATGAACACCC	4636
QY	1562	CGGGAATTCCTCGGTGCGCAAGCACTCTTTGAATTTTGGAGGGCGCTTTTACCGGTCTCA	1621
DB	4637	CGGGAATTCCTCGGTGCGCAAGCACTCTTTGAATTTTGGAGGGCGCTTTTACCGGTCTCA	4696
QY	1622	CCCATATACCGCCACTTCTCTATCCAGACAAGCAGAGTGGGGAACCTTCCCTATC	1681
DB	4697	CTCATATATGCGCCACTTCTCTATCCAGACAAGCAGAGTGGGGAACCTTCTCTTACC	4756

QY 781 ACTACTGGCAGCCCGATCAGTATTCACCTACGGCAAGTTCCTTCCGACCGCGGGTGT 840
Db |||||
QY 781 ACCACTGGCAGCCCGATCAGTATTCACCTACGGCAAGTTCCTTCCGACCGCGGGTGT 840
Db |||||
QY 841 TCAGGGGGTCTTATGACATAAATAATTTGTGACGAGTGCCACTCCACGGATGCAACATCC 900
Db |||||
QY 841 TCAGGGGGGGCTTATGACATAAATAATTTGTGACGAGTGCCACTCCACGGATGCAACATCC 900
Db |||||
QY 901 ATCTTGGGCAATTCGGCACTGTCCTTGAACAGAGAGACCGCGGGGCGAGACTGACTGTG 960
Db |||||
QY 901 ATCTTGGGCAATTCGGCACTGTCCTTGAACAGAGAGACTGCGGGGCGAGACTGCTGTG 960
Db |||||
QY 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAACATCGAGGAGGTT 1020
Db |||||
QY 961 CTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAACATCGAGGAGGTT 1020
Db |||||
QY 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTGNAGCAAT 1080
Db |||||
QY 1021 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATTCCTTTCGAAGTAATC 1080
Db |||||
QY 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTGCGACGAGCTCGCGCGCA 1140
Db |||||
QY 1081 AAGGGGGGAGACATCTCATCTTCTGTCTTCAAAGAGAGTGCGACGAGCTCGCGCGCA 1140
Db |||||
QY 1141 AAACCTGGTCGGTGGGGCGTCAATGCGGTGGCTTACTACCGGGGCGCTTGATGTCCGTG 1200
Db |||||
QY 1141 AAGCTGGTCGCTTGGGCATCAATGCGGTGGCTTACTACCGGGGCTTGACGCTGTCGGTC 1200
Db |||||
QY 1201 ATCCGACGAGTGGTGAAGTGTGCTGTGGGCAACTGACGCCCTCATGACCGGCTTACC 1260
Db |||||
QY 1201 ATCCCGCCCATCGGCGATGTTGTGCTGTGGCAACCGATGCCCCCTCATGACCGGCTATACC 1260
Db |||||
QY 1261 GGGGACTTCGATTCGGTGATAGACTGCAACACGCTGTGTCACCCAGACAGTTCGACTTTCAGC 1320
Db |||||
QY 1261 GGGGACTTCGACTCGGTGATAGACTGCAATACGCTGTGTCACCCAGACAGTTCGACTTTCAGC 1320
Db |||||
QY 1321 CTTGACCCCTACCTTACCATTGAGACAATCAGCTTCCCGAGATGCTGTCTCCCGTACT 1380
Db |||||
QY 1321 CTTGACCCCTACCTTACCATTGAGACAATCAGCTTCCCGCAAGATGCTGTCTCCCGCACT 1380
Db |||||
QY 1381 CAACGTGGGGTAGGACTGGCAGAGGAGGAGCCAGGATCTACAGATTTGTGGACCCGGGG 1440
Db |||||
QY 1381 CAACGTGGGGGAGGACTGGCAGGGGGAAGCCAGGATCTACAGATTTGTGGCACCAGGG 1440
Db |||||
QY 1441 GAGCGCTCTTCTGCGCATGTTGACTCGTCTGCTCTCTGGAGTGCTATGACGCGGCTGT 1500
Db |||||
QY 1441 GAGCGCCCTCCGGCATGTTGCACTCGTCCGTCTCTGTGAGTGCTATGACGAGGCTGT 1500
Db |||||
QY 1501 GCTTGTGTAGCTTACGCCCGCGAGACCAAGTTAGGCTACGAGCATACATGAACACC 1560
Db |||||
QY 1501 GCTTGTGTAGCTTACGCCCGCGAGACTACAGTTAGGCTACGAGCATACATGAACACC 1560
Db |||||
QY 1561 CCGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGTCTTACGGGTCTC 1620
Db |||||
QY 1561 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGTCTTACAGGCCTC 1620
Db |||||
QY 1621 ACCACATAGAGCCCACTTCTCTATCCAGCAAAAGCAGAGTGCGGGAACCTTCCCTAT 1680
Db |||||
QY 1621 ACTCATATAGATGCCCATCTTCTATCCAGCAAAAGCAGAGTGCGGGAACCTTCTTAC 1680
Db |||||
QY 1681 CTGGTAGCGTACCAAGCACCGTGTGCGCTAGAGCTCAAGCCCCCTCCCGCTGTGGGAC 1740
Db |||||
QY 1681 CTGGTAGCGTACCAAGCACCGTGTGCGCTAGGCTCAAGCCCCCTCCCGCTGTGGGAC 1740
Db |||||
QY 1741 CAGATGTGAAGTGTGATCCGTCTCAAGCCCAACCTTCATGGGCGCAACCTCTGTGCTA 1800
Db |||||
QY 1741 CAGATGTGAAGTGTGATTCGCTCAAGCCCAACCTTCATGGGCGCAACCTCTGTGCTA 1800
Db |||||
QY 1801 TATAGACTGGGCGCTGTCCAGATGAGTCACTGACCGCACCGAGTCAACAGTATATC 1860
Db |||||
QY 1801 TACAGACTGGGCGCTGTTCAGATGAATCACTGACCGCACCGAGTCAACAAATATATC 1860
Db |||||

QY 1861 ATGACATGATATCGGCTGACCTGGAGGTGCTCAAGAGTACCTGGGTGCTCGTTGGCGGC 1920
Db |||||
QY 1861 ATGACATGATATCGGCTGACCTGGAGGTGCTCAAGAGTACCTGGGTGCTCGTTGGCGGC 1920
Db |||||
QY 1921 GTTCTGGCTGCTTGGCCGCGTATTGCCCTATCCACAGGCTGCTGTCATAGTAGGTAGG 1980
Db |||||
QY 1921 GTTCTGGCTGCTTGGCCGCGTATTGCCCTGTCAACAGGCTGCTGTCATAGTAGGTAGG 1980
Db |||||
QY 1981 ATTGTCTTGTCCGGAAAGCCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTTC 2040
Db |||||
QY 1981 GTCTGTCTTGTCCGGGAGCCGGCAATCATACCTGACAGGGAAGTCTCTACCGGAGTTTC 2040
Db |||||
QY 2041 GATGAAATGGAAGAGTGC 2058
Db |||||
QY 2041 GATGAGATGGAAGAGTGC 2058
Db |||||

Search completed: February 27, 2005, 22:43:12
Job time : 376 secs

THIS PAGE BLANK (USPTO)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C	1	46.8	2.3	509	9	CL252974
	2	44.8	2.2	935	9	CNS006XK
	3	44.4	2.2	834	4	BI956973
	4	43.4	2.1	666	5	BM950570
	5	42.8	2.1	1863	7	CV068974
	6	42.6	2.1	633	4	BI959933
	7	42.2	2.0	871	6	CD437613
	8	41.8	2.0	564	2	BE291962
	9	41.8	2.0	673	4	BI155705
	10	41.8	2.0	935	5	BQ876236
	11	41.8	2.0	1020	6	BY117119
	12	41.8	2.0	1409	3	AK012576
	13	40.8	2.0	925	9	CNS0091P
	14	40.4	2.0	535	4	BJ277484
	15	40.4	2.0	1749	9	CL948408
	16	40.2	2.0	574	4	BJ208789
	17	40.2	2.0	629	4	BG906349
	18	40.2	2.0	684	4	BI956133
	19	40.0	1.9	697	7	CN944784
	20	39.8	1.9	624	6	CD938118
	21	39.8	1.9	896	4	BQ969213
	22	39.6	1.9	424	5	BQ65971
	23	39.6	1.9	429	5	BQ665888
	24	39.6	1.9	430	5	BQ665954


```

/NOCC= Organ: brain; vector: p1x abc; p1c1: 1.000 1;
Site 2: Not I; The library was constructed according to

```

Query Match 2.1%; Score 42.8; DB 7; Length 1863;
Best Local Similarity 48.4%; Pred. No. 3.7;
Matches 119; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
QY 1126 GACGAGCTCCGCCAAAACGTGTCGCGTGGCGGCTCAATGCCGTGCTTACTACCGGGC 1185

Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see close TJ, Wing R, Kleinborts A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. <http://www.clemson.edu/barl/31/cover.html>

ORIGIN

Query Match	2.1k	Score	42.6	DB	4	Length	633
Best Local Similarity	51.3	Pred. No.	3.1	94	Indels	0	Gaps
Matches	99	Conservative	0	Mismatches	0	Indels	0
1096	CTCATCTTCGCCACTCAAGAAGAACTGGCGAGAGCTCCCGCAAACTGGTGCCTTG	1155					
405	CTCTTCAACTGCTCCAGCGGACAGACTGGCGGCGCCCAACGGAGCGGTTTCGCGCG	464					
1156	GGCGTCAATGCGGCGTTCATCCGGCGCCTTGATGTGTCGGTCACTCCCGACCACTGGT	1215					
465	AGCATCAACACACGCTCTTCGTCTCCCTCCACCGTCTCCATCTTCCAGGCACATAC	524					
1216	GACGTTGTCGTCGCGCAACTGACGCGCCTCATGACCGGCTTTACCGCGCACTTCGATTCG	1275					
525	CAGGGGATGCCGGCGGCAACGGGCTCTTTCACCAACCGGCTTTCCCGCGCAACCCCGCGTG	584					
1276	GTGATAGACTGCA	1288					
585	CAGTTTCGACTACA	597					

RESULT 7

	CD437613	871 bp	mRNA	linear	EST 03-JUN-2003
	CD437613	5 Zea mays cDNA.	mRNA sequence.		
	LOCUS				

DEFINITION
ACCESSION

ACCESSION
VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

REFERENCES

.....

TITLE

1

JOURNAL
COMPTON

COMMENT

DECLASSIFIED

FEATURES

0400

ORIGIN

[illegible][illegible]

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
BI959933		EST.	Hordeum vulgare subsp. vulgare	
BI959933_1	GI:16311188		Hordeum vulgare subsp. vulgare	
			Eukaryota; Viridiplantae; Strestophyta; Embryophyta; Tracheophyta;	
			Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
			Proideae; Triticeae; Hordeum.	

REFERENCE
AUTHORS
1. Wing, R., Close, T.J., Kleinhofes, A., Wise, R., Chin, A., Begum, D., Prisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Oates, R. and Main, D.
TITLE
Development of a genetically and physically anchored EST resource for barley genomics: Morex rachis cDNA library
JOURNAL
Unpublished (2001)
COMMENT
Contact: Wing RA

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 854 556 7288
Fax: 854 556 4293
Email: rwings@clemson.edu
Total hg bases = 513
Seq primer: [RAATTAACCCCTCACTAAAGGG](#)

```

FEATURES
source
High quality sequence status: 595.
Location/Qualifiers
1..633
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="rRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"

```

```

/clone="HVSME0022H07f"
/tissue.type="rachis"
/lab.host="rJCL21"
/clone.lib="Hordeum vulgare rachis EST library HVCDA0015
(normal)"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Plants were grown at Washington State University,
Pullman, WA in a greenhouse, the rachises were excised and

```

frozen in liquid nitrogen (Kleinfuchs lab). In the TJ Clonson lab at the University of California, Riverside total RNA was prepared, poly(A)⁺ was purified, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids (Chin). Phagemids were plated and picked at the Clonson University Genomics Institute (CUGI) (Begum, Palmer,

```
Query Match      2.0%; Score 42.2; DB 6; Length 871;
Best Local Similarity 48.9%; Pred. No. 4.4;
Matches 113; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1144 CTGGTCGGTGGCGCTCAATGCGGTGGCTTACTACCGCGGCTTGTATGTCCTGCTCATC 1203
DB 151 CTGCGCGCGAGTGGCTCGCGGACATTTGTCGAGTACCTCGCGACCGCTTTCTCTCGTCGAC 210
QY 1204 CCGACCAAGTGGTGAAGTGTGCTGGTGGCACTGAGCCCTCATGACCGGCTTTACCGGC 1263
DB 211 CCGGCTTCTCGCGACATCGAGTCACTGAGTCAAGGACCCCTACCGGC 270
QY 1264 GACTTCGATTTGGTGTAGTCAACACAGTGTGTCAACCCAGAGTGTGACTTTCAGCCTT 1323
DB 271 GAGCTCGACGGCGGCTCTCGCGCGGCTCGCTCGGTATCTCGAGGCGGAGTTCCGCGCG 330
QY 1324 GACCTACCTTACCATTTGAGACATCACGCTTCCCGAGATGCTGTCTCC 1374
DB 331 CTCTCGCCGACCACTCCGCGCGCTCGGATGCGGACCGCGCGCGCC 381

RESULT 8
BE291962/c
LOCUS      601085590F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3499726 5',
DEFINITION mRNA sequence.
ACCESSION BE291962
VERSION    BE291962.1 GI:9174040
KEYWORDS
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus

REFERENCE
1 (bases 1 to 564)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8557 row: c column: 23
High quality sequence stop: 561.
Location/Qualifiers
1. .564
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3499726"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam6"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

ORIGIN
Query Match      2.0%; Score 41.8; DB 2; Length 564;
Best Local Similarity 53.3%; Pred. No. 5;
Matches 88; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 687 CAACCCCTCGTGTGCAACAATGGGCTTTGGTGTATCATGTCCAAGGCCCATGGGAT 746
|||||
|||||

Query Match      2.0%; Score 41.8; DB 4; Length 673;
Best Local Similarity 53.3%; Pred. No. 5.3;
Matches 88; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 687 CAACCCCTCGTGTGCAACAATGGGCTTTGGTGTATCATGTCCAAGGCCCATGGGAT 746
|||||
|||||
397 CAAGGCCACCTGTCTGGAGCCATGGGCTCTCGAGTGTCTCTCGCTCAGAGCCACAGGGCT 338
|||||
|||||
747 TGATCTTAACATCAGACCTGGGGTGAGGACAATTAATCTACTGGAGCCCATCAGTATTC 806
|||||
|||||
337 GGACACAGAGGGCAGCTCTCTCGGGTAGGCAAGGTCACAGGAGCCTGGGGCTAGTGCC 278
|||||
|||||
807 CACCTACGCAAGTTCCTTTGCCGACGGGGGTTTCAGGGGGTGC 851
|||||
|||||
277 CGGCTCTGGAGCAGCTGCGAGCCCGCGGTGGAGTTTCAGCTGAGGC 233
|||||
|||||

Db 266 CAAGGCCACCTGTGTGGAGCCATGGGCTCTCGAGTGTCTCGCTCAGAGCCACAGGGCT 207
747 TGATCTTAACATCAGACCTGGGGTGAGGACAATTAATCTACTGGAGCCCATCAGTATTC 806
206 GGACACAGAGGGCAGCTCTCTCGGGTAGGCAAGTCCACAGGAGCCTGGGGCTAGTGCC 147
807 CACCTACGCAAGTTCCTTTGCCGACGGGGGTTTCAGGGGGTGC 851
146 CGGCTCTGGAGCAGCTGCGAGCCCGCGGTGGAGTTTCAGCTGAGGC 102

RESULT 9
BI155705/c
LOCUS      602904453F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5033964 5',
DEFINITION mRNA sequence.
ACCESSION BI155705
VERSION    BI155705.1 GI:14615706
KEYWORDS
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus

REFERENCE
1 (bases 1 to 673)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1094 row: j column: 13
High quality sequence stop: 627.
Location/Qualifiers
1. .673
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5033964"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam3"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

ORIGIN
Query Match      2.0%; Score 41.8; DB 4; Length 673;
Best Local Similarity 53.3%; Pred. No. 5.3;
Matches 88; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 687 CAACCCCTCGTGTGCAACAATGGGCTTTGGTGTATCATGTCCAAGGCCCATGGGAT 746
|||||
|||||
397 CAAGGCCACCTGTCTGGAGCCATGGGCTCTCGAGTGTCTCTCGCTCAGAGCCACAGGGCT 338
|||||
|||||
747 TGATCTTAACATCAGACCTGGGGTGAGGACAATTAATCTACTGGAGCCCATCAGTATTC 806
|||||
|||||
337 GGACACAGAGGGCAGCTCTCTCGGGTAGGCAAGGTCACAGGAGCCTGGGGCTAGTGCC 278
|||||
|||||
807 CACCTACGCAAGTTCCTTTGCCGACGGGGGTTTCAGGGGGTGC 851
|||||
|||||
277 CGGCTCTGGAGCAGCTGCGAGCCCGCGGTGGAGTTTCAGCTGAGGC 233
|||||
|||||
```


Nature	420,	563-573	(2002)
6	(bases 1 to 1409)		
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hirooka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,C.H., Kawai,J., Kojima,Y., Komoto,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu.M. and Hayashizaki, Y.			
Submitted	(10-JUL-2000)		
Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)			
Please visit our web site	(http://genome.gsc.riken.jp/) for further details.		
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGATCCACAGCTCTTTTITTTTTTIVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCTCGAGTTAATAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.			
Host:	SouR.		
Location/Qualifiers			
1..1409	/organism="Mus musculus"		
/mol_type="mRNA"			
/strain="C57BL/6J"			
/db_xref="FANTOM_DB:2700087H15"			
/db_xref="taxon:10090"			
/clone="2700087H15"			
/tissue_type="whole body"			
/clone_lib="RIKEN full-length enriched mouse cDNA library"			
/dev_stage="11 days embryo"			
566..1360	/note="unnamed protein product; hypothetical protein (evidence: NCBI CDS Predictor,Longest-ORF) putative"		
/codon_start=1			
/protein_id="BAB28329.1"			
/db_xref="GI:12849409"			
/translation="MQACVMPSSLQNNMGFSSEVARQTDPREPPAPPOATPWP SQGSWARRGFPAWSQPALSOOKLRRCLGLGDSSRTSLICPTLHRPALSPQ RGCGLPPLSVGLWGTEPYTSNGRPFSALGCHFRKSRTRYPCNLTPALPTLCRGGA SASOARADGLLLPSCPSPFWGLGHSSGSTPTCPGVGLVSLEAPSVAPIQSASL ILFLMVLFTMSVGPEQAAPPFGVRAVLFEPRLDSTVPREDTVM"			
1390..1395	/note="putative"		
polyA_signal			
1409	/note="putative"		
polyA_site			
ORIGIN			
Query Match... 2.0%; Score 41.8; DB 3; Length 1409;			
Best Local Similarity 53.3%; Pred. No. 6.4;			
Matches 88; Conservative 0; Mismatches 77; Indels 0; Gaps 0;			
QY 687 CAACCCCTCGTGTCGCAACAATGGCTTTGTGCTTACATGTCACAGGCCCATGGAT 746			
Db 352 CAAGGCCACTGCTGTGGAGCATGGGCTCTGCAGTGTCTCGCTCAGACCAGGGCT 293			
QY 747 TGATCTTAACATCAGGACTGGGGTGAGGACAAATTACTAGCGCACCGCATCATTC 806			
Db 292 GGACACAAGGGCAGCTCTCGGGTAGGCAAGGTCCACAGGACGCTGGGGCTAGTGCC 233			
QY 807 CACTTAGCGCAAGTTCCTTTCGCGAGCGGGGTGTCAGGGGGTGC 851			
Db 232 CGGCTCTGGAGCAGCTGTCAGCGCGCGGTGGGAGTTTCAGCTGAGGC 188			
RESULT 12			
AK012576/c			
LOCUS			
DEFINITION			
AK012576 Mus musculus 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2700087H15 product:hypothetical protein, full insert sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
Carninci,P. and Hayashizaki, Y.			
High-efficiency full-length cDNA cloning			
Meth. Enzymol. 303, 19-44 (1999)			
99279253			
10349636			
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki, Y.			
Normalization and subtraction of cap-trapper-selected cDNAs to Prepare full-length cDNA libraries for rapid discovery of new genes			
Genome Res. 10 (10), 1617-1630 (2000)			
20499374			
11042159			
Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunaga,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki, Y.			
RIKEN integrated sequence analysis (KISA) system--384-format sequencing pipeline with 384 multicapillary sequencer			
Genome Res. 10 (11), 1757-1771 (2000)			
20530913			
11076861			
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
Functional annotation of a full-length mouse cDNA collection			
Nature 409, 685-690 (2001)			
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.			
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs			

747 TGATCTTAACATCAGGACTGGGTGAGGACAATTACTACTGGCAGCCCATCACGTTATTC 806
Db GGACACAAGGGGCAGTCTCTCGGTATGGCAAGTCCACAGGACGCTGGGGCTAGTGCC 233

QY 807 CACTACGGCAAGTTCCTTCCGCCACGCGGTGTTCAGGGGGTGC 851
Db CGGCTGGAGCAGCTGCAGCCCGGTGGAGTTTANGCTGAGGC 188

RESULT 13
CNS0091P 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR19D16 of RPCI-98 library from drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION AL053013
KEYWORDS AL053013 GI:4934461
SOURCE GSS.
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Epiphyarioidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequences :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr
COMMENT - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammossier in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source 1..925
/organism="Drosophila melanogaster"
/mol_type="Genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 2.0%; Score 40.8; DB 9; Length 925;
Best Local Similarity 12.9%; Pred. No. 11;
Matches 47; Conservative 159; Mismatches 158; Indels 0; Gaps 0

QY 346 CTGTGTCGCGACGGGTGATGGCAGGGGACCGCTGTTTCGCCCGCGCTATCTCTTAC 405
Db SCSSBSGSSSNMTSSNSSBCSCSSBSSTSSMSSSSSSSGSSSGTSACV 857

QY 406 TTGAAGAAGGCTCTCGGAGGCCCTCTGTGTGCCCGCCAGGACATCCGTAGGCATATTC 465
Db KCVASSSCCGCGCMABCCWCSSSSCCSASARGVKVRASGAGKRGGGGSASHS 797

QY 466 AGAGCCCGGTATGCACCCTGGAGTGGCTAAGGCGGTGCATCTCATCCCCGTAGAGCC 525
Db SSACBSSSSSCASWSASSSASSSRSRRSGGAGGSSASSRSSSSSSASAGSVSS 737

QY 526 TTAGACACAACCATGAGGTCGCCGGTGTCTTCAGACAACATCTCTCCCACACAGAGTGCCC 585
pb ASSSSSSSSSVSCSVASMSGSSSSSSSSASSSSSSSSSASCSACSCCTSWGCSCT 677

Search completed: February 27, 2005, 22:36:51
Job time : 6109 secs

THIS PAGE BLANK (USPTO)